DOCUMENT RESUME

ED 242 268 HE 017 133

TITLE The Target Income Hypothesis and Related Issues in Health Manpower Policy.
INSTITUTION Health Resources Administration (DHEW/PHS), Bethesda, Md. Bureau of Health Manpower.
REPORT NO HRA-80-27
PUB DATE Jan 80
NOTE 105p.; For related documents, see HE 017 122-124 and HE 017 127-132.
PUB TYPE Collected Works - General (020) -- Viewpoints (120) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS *Costs; *Dentists; Economic Factors; Employment Patterns; *Health Services; Higher Education; *Income; *Labor Supply; *Physicians; Public Policy; Theories

ABSTRACT

The target income hypothesis as applied to health personnel is discussed in seven papers and three commentaries. The hypothesis pertains to the economic behavior of physicians, and it has recently been applied to dentists as well. The issues concern the production/supply of practitioners and the resulting cost for their services. Whether physicians/dentists can induce demand for their services and set income in accordance with a preconceived target has important policy implications. Paper titles and authors are as follows: "Introduction" (Jesse S. Hixson); "An Analysis of Competing Hypotheses of the Demand for and Supply of Physician Services" (James B. Ramsey); "A Formal Model of Target Income Pricing with Supplier-Induced Demand" (Charles S. Roehrig); "The Effect of Provider Supply on Price" (Mark V. Pauly, Mark A. Satterthwaite); "The Aggregate Supplies and Demands of Physician and Dental Services" (Jesse S. Hixson, Nina Mocniak); "The Effect of Local Physician Supply on the Treatment of Hypertension in Quebec" (Philip J. Held, Larry M. Manheim); "A Market Model of the Distribution of Dentists" (Gerald L. Musgrave); "Conflicting Theories of the Determination of the Entrepreneur's Income: An Analysis of the Practicing Dentist" (Donald R. House); and discussions by Frank A. Sloan, Zachary Y. Dyckman, and Thomas R. Saving. (SW)

D. For each given value of D the analysis is the same as that shown in Figure 1. For notational convenience, the subscript "f" on x_{nf} can be dropped without confusion, since the ensuing analysis need deal only with finding the optimal value of π.

Letting $H_1(x_n,D)$, $H_2(x_n,D)$ denote the normal equations obtained from the first order conditions, one

The above formulation enables one to compare the theoretical predictions of the Evans model to those of the corresponding standard model; in the current formulation the standard model results are obtained by setting all partial derivatives with respect to D to zero.

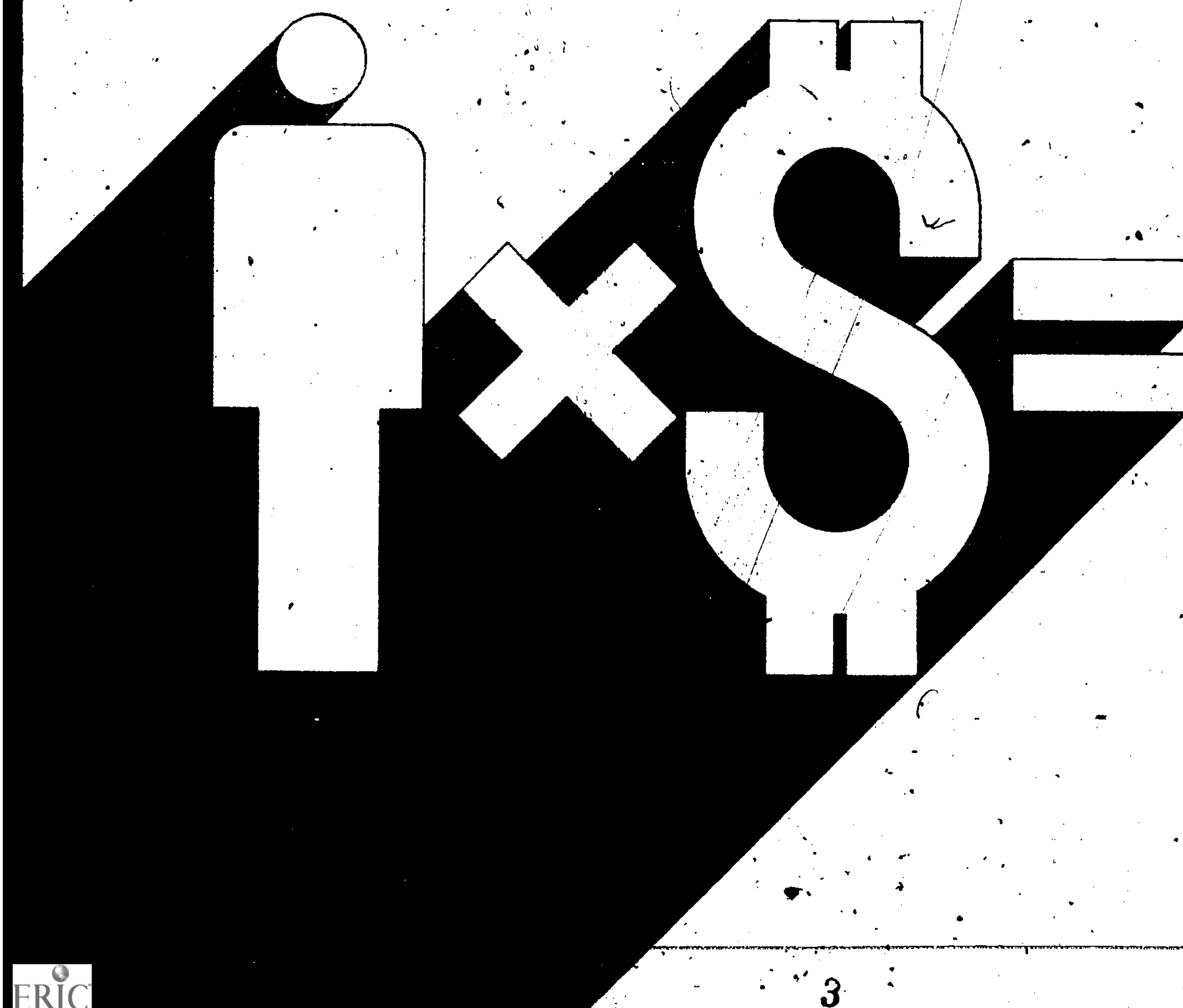
Consider an increase in w_n, the firm's opportunity cost for physician input. Since both H_{1w}, H_{2w} = (- 1),

The Target Income Hypothesis

And Related Issues in Health Manpower Policy

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE
Public Health Service
Bureau of Health Manpower
Division of Manpower Analysis
and
Division of Dentistry

DHEW Publication No. (HRA) 80-27



Contents

Introduction

Jesse S. Hixson
Bureau of Health Manpower
Health Resources Administration

How many physicians and dentists does the Nation need? What should the Federal Government do to improve the geographical distribution of the Nation's physicians and dentists? These questions come to the forefront each time the Congress debates the renewal or revision of its legislation for financial assistance to health professional training institutions and students.

Historically, the debate has been heavily flavored with rhetoric about shortages of health manpower relative to unmet health care needs of the population. Recently, however, the debate has taken a new twist. Many are repeating the argument that the Nation has reached the point where continued production of physicians and dentists will result only in an increase in prices of health care, not in the amount of care produced and consumed. The implication of this argument is twofold: the Nation's output of physicians and dentists should be decelerated, and drastic Federal intervention is required to assure that physicians and dentists are distributed to provide optimally the health services that the Nation demands.

The basis of this argument is the so-called "target income hypothesis" about physicians' and dentists' pricing behavior. The target income hypothesis has developed a significant following among economists doing empirical research on physicians' economic behavior, and it has recently been pursued in empirical work on dentists' supply behavior as well. Despite the attention it has received, the origin of the "hypothesis" remains obscure. It has never been formally stated, and its empirical implications have never been formally derived or tested. It is most often encountered as a blatantly naive assertion. As an example, the following is a passage from a document prepared for the Council of State Governments discussing current issues affecting dentistry. Referring to what Robert G. Evans has termed the "utilization promotion effect" in his empirical study of physician incomes [2], the document states:

> The same theory can be applied to dentistry. If correct, continued production of dental graduates, thought by many to be the source of more dental care for more people, will probably result in higher costs for dental services. For example, assuming the present demand for dental services and the concept of target income, a simple equation related to the number of dentists can be developed:
>
> number of dentists x target income = total expenditure for dental care
>
> This indicates that an increase in the number of dentists, in their target income, or in both will result in increased dental expenditures, regardless of the number of consumers . . . . [1]

The document goes on to suggest that to protect the public from the insidious consequences of a continuing output of dentists, the Government should institute demand-augmenting programs. The document recommends that to protect the public from the alleged perverse consequences of "too many" dentists, the Government should take measures that will result in enhancing the dental profession's economic status. Thus the document makes use of the "target income" hypothesis to argue that it is really in the public interest to pursue a policy that, from the perspective of orthodox economic theory, could be construed to be solely in the interest of the dental profession and detrimental to the public interest.

Although the origin of the target income hypothesis is obscure, one can point to three papers published in the early 1970's that precipitated a major interest in the notion and convinced many economists that orthodox economic models do not apply to the health care markets. Two of the papers were published in 1970 —

these were Feldstein's "The Rising Price of Physicians' Services" [4] and Newhouse's "A Model of Physician Pricing" [5] — while the third was the 1973 article by Evans, Parish, and Sully, "Medical Productivity, Scale Effects, and Demand Creation." [3] Since these papers are reviewed elsewhere in this volume, only a few brief comments about their apparent impact on the subsequent literature are made here.

None of the papers presents a well-defined concept of target-income pricing behavior or a formalization of the hypothesis from which empirical implications can be derived. Thus, none of the authors has a theory on which to base the interpretations of his empirical results as representing income-targeting behavior. The empirical results of the first 2 papers are derived from 18 to 20 observations. Those of the first paper are presented as total failures to fit conventional models to aggregate time-series data, with the conclusions presented as speculative rationalizations of the unsuccessful search for the expected relationships. The conclusions of the second paper are based on samples of up to 20 observations in which assorted "demand variables" included in price equations are never significant and often have the "wrong" sign. The author suggests that the results may be due to the high correlations between the variables and the low number of observations, but prefers the speculative conclusion that physicians are not "fully" profit maximizers but satisfiers with income targets! The conclusion in favor of target income behavior presented in the third paper rests on the authors' demonstration that output per physician is inversely related to local physician/population ratios. Despite their pointing out several crucial limitations of their "test" for supply-induced demand, they preferred to emphasize the "consistency" of the result with the demand-creation hypothesis.

One can only say that the results of these three studies are inconclusive at best. Nevertheless, these three papers are often cited as strong evidence of unorthodox pricing behvaior in the market for health services. It has become fashionable to include a practitioner-to-population ratio variable in almost every investigation of physician or dentist service supply or demand to capture the expected demand-creation phenomenon. A recent study of the dental service input-output relationship even included the dentist-population ratio in the estimated production function to test for supplier-induced demand [6].

At the present time, the debate about the merits and relevance of the target income hypothesis and related issues is far from resolved among economists. In fact, it has not progressed to the point where the meaning of the terms used by the protagonists and antagonists are agreed upon by all those engaged in the debate. Nevertheless, the presumed implications of the hypothesis are being contemplated by policymakers and advisors to policymakers, and are having profound effects on their thinking and their recommendations for health and health manpower policy. Despite the flimsiness of the theoretical basis and the empirical support for a belief that the phenomena exist, "target income" behavior and "supplier-induced demand" seem to provide an attractive rationale for the abruptly changing public attitude toward support of health professions education and toward regulation of behavior in the health care system.

Although the politics of the situation may have predetermined the outcome of the public policy debate, it is not too late to examine the logical basis of the theories and the merits of the empirical evidence, as well as to attempt to develop a rigorous approach to resolving the issues in the scientific dimension. Toward this end, the Bureau of Health Manpower has asked a number of individuals, most of whom have not been involved in the target-income controversy, to view the issues from various perspectives. Their papers, together with comments from invited discussants, are included in this volume which, it is hoped, will go some way toward providing more resolution to the issues.

References

1. The Council of State Governments, *Manpower Utilization: A Background Paper Prepared for the National Task Force on State Dental Policies* (Mimeographed) 1978
2. Evans RG: Does Canada have too many doctors? Why nobody loves an immigrant physician. *Canada Public Policy — Analyses Politiques* 2 (Spring, 1977)
3. Evans RG, Parish EMA, Sully F: Medical productivity, scale effects, and demand generation. *Canadian Journal of Economics* VI: August 1973
4. Feldstein MS: The rising price of physicians' services. *The Review of Economics and Statistics* LII: May 1970
5. Newhouse JP: A model of physician pricing. *Southern Economic Journal* 37: October 1970
6. Sheffler RM, Kushman JE: A production function for dental services: Estimation and economic implications. *Southern Economic Journal* 44: July 1977

An Analysis of Competing Hypotheses of the Demand for and Supply of Physician Services

James B. Ramsey, Ph.D.
New York University

ABSTRACT ☐ A theoretical framework is provided in order to analyze the implications of competing hypotheses about the economic behavior of physician firms; special attention is paid to the so-called "supply-induced demand" and target income hypotheses. In this connection feasible hypothesis test procedures are developed in order to discriminate between "supply-induced demand" and the more traditional models of physician firms.

The second objective of the paper is to indicate the role specification error tests could play in evaluating the empirical relevance of the various hypotheses.

The final objective of the paper is to indicate how standard theory might be applicable to the observed data, and how the prior empirical analysis should be re-examined in order to separate theoretically important findings from statistical artifacts.

INTRODUCTION

Over the last eight years a considerable controversy has been going on concerning the appropriate theoretical models needed to explain demand and supply relations in the physician and dental services market. The controversy arose after several researchers discovered empirical results that were at variance with the simple formulations of standard economic theory used in the analysis. The main statistical facts causing trouble were the observation of positive partial correlations between physician/population ratios and physician fees and the discovery of positive price elasticities of demand. This issue is still of great interest to policy makers as witnessed, for example, by the discussion in [6].

A number of ad hoc rationalizations were suggested, but two vaguely formulated competing hypotheses in particular were proposed. The first was that physicians could alter the demand for their services directly by persuading their patients to consume more, or less, at current prices. The second hypothesis was that doctors acquired in some unspecified manner or inherited some notion of an "ideal" or target income. They would reduce their efforts if actual income rose above the target and increase them if actual income fell below. Both of these hypotheses went through numerous changes and modifications.

The objectives of this paper are quite straightforward. The first step is to provide a theoretical framework within which to analyze the theoretical implications of the competing hypotheses, evaluate the differences between them, and explore the extent to which empirical observations can be used to discriminate between the competing hypotheses.

The second objective is to indicate the extent to which specification error tests could have been used to evaluate the inferences drawn from the previous research. If the models are seriously misspecified, no useful inferences can be drawn from the observed statistical relationships.

The third task is to indicate some of the major ways of reinterpreting the observed data so that versions of the standard theory may be applicable. In short, this part of the paper leads one to suspect, in addition to the usual *caveats* for the applied economist, that the models used so far in the literature have been seriously misspecified.

The first part of the paper briefly reviews the extant literature insofar as it bears on the objectives of this paper. A much more thorough and elaborate review is found in Sloan and Feldman [38]. The second part, the heart of the paper, is theoretical and anlyzes the theoretical structures of the alternative models in detail. The third section is concerned with the prevalence of specification error in the models used in prior research and indicates the need for specification error tests.

PART I: A SCHEMATIC REVIEW OF THE EMPIRICAL LITERATURE

Three board areas of concern to researchers have been evidenced in the literature over the past 8 years. One of the earliest areas of empirical dispute was whether the physician maximized his personal utility or profit and, if the latter, was he a price taker or a price searcher? The second area of interest concerned the fact that many researchers obtained empirical results inconsistent with the use of simple versions of standard microeconomic theory in the physician market; some of these results were in terms of time series, others in terms of cross-sectional studies. The third category is a miscellany of work on physician firm production functions, migration of physicians, and quality variations.

A comprehensive, detailed, and recent discussion of the various empirical findings is contained in Sloan and Feldman [38]. The discussion presented here is meant as a broad guide to the principal participants in the various areas under discussion and a recapitulation of the major findings; the theoretical models and the econometric procedures are not questioned in detail. Needless to say, each and every study under discussion suffers more or less from the theoretical criticisms of current models developed in Part II, the problems caused by lack of identification, inappropriate use of proxy variables, and from a lack of examination of the estimated models for specification errors.

Empirical Evidence on Utility vs. Profit Maximization, Price-taking vs. Price-searching Behavior

The most formal and specific debate on the issue of whether physician firms are monopolistically competitive (price searchers) or competitive (price takers) was stimulated by Newhouse's attempt in [21] to discriminate between the two hypotheses. Unfortunately, his formulization of the problem was incorrect, as pointed out by French and Ginsburg [15]. However, later work by Kushman and Scheffler [20] on dentistry appears to discriminate between price-searching and price-taking hypotheses in favor of the former. While most researchers who assume profit-maximizing behavior assume the physician firm is a price searcher, the issue is by no means settled. One of the difficulties is that while there may be a form of price-taking behavior in this market, the presence of entry barriers and the state of consumer and physician ignorance of the market because of the lack of advertising contribute to a wide variation in price, as shown in Newhouse and Sloan [22]. The result is that tests to discriminate between *stochastic* formulations of the two hypotheses become difficult. Holen [18] has examined the effects of dental licensing on the "quality of dental care" and concluded that quality was improved. The author did not examine other market effects.

While many researchers assume that physicians are price-searching profit maximizers, everyone else assumes that physicians are utility maximizers who trade profit and income for leisure, interesting cases, target incomes, and whatever else strikes a researcher's fancy. There have been no formal tests of the hypothesis of profit maximization as opposed to utility maximization. However, the apparent rejections of some implications of simple models of profit maximization have been interpreted as evidence in favor of the utility-maximizing approach.

Empirical Evidence and the Demand for and Supply of Physician Services

The first and relatively sophisiticated study of the physician market was based on time-series data, 1948-1966, and was carried out by Feldstein [11]. Feldstein was one of the first researchers to make extensive use of the distinction between the average price (fee) received by the physician and the net price (fee) paid by the insured patient. Feldstein postulated some plausible relationships between what he regarded as relevant economic variables and examined in turn a simultaneous equation system and a dynamic adjustment system. The "equilibrium demand equation" related quantity of physician services per capita to median income, per capita provision of government services, net or average price, the C.P.I., and a time trend. Both the price variables and the insurance variable were found to be statistically significant; but with "incorrect signs" on the basis of a standard model. In the "dynamic" version of the model a "reduced form" equation was derived for the logarithm of average price. Further, by assuming that the observations on quantity traded in the market identified the supply equation, Feldstein obtained the empirical results that quantity supplied was negatively related to price and demand positively related. While the equations may be intuitively appealing, they are hard to interpret in terms of the structural equations one would obtain from utility and profit-maximizing behavior; for example, in the supply equation a variable called "reference income" is used as "an attempt to capture the effect of physicians' rising income aspirations (i.e., the higher "marginal utility" of income)" [11, p.131].

In any event, the two most important conclusions Feldstein drew from his study were that there is a per-

manent excess demand for physician services and that physicians reduce output in reaction to a rise in fees [12]. Brown and Lapan [4] criticized the Feldstein results on the basis of his definition of price and the potential identification problem in the Feldstein equations. The best that can be said about the resulting debate with Feldstein is that the empirical results are somewhat sensitive to obvious changes in the functional form of the equations and to alternative definitions of price.

One of the other early major studies to apply an apparently sophisticated model (a simultaneous equation system was formulated) to the data was that by Fuchs and Kramer [16]. In the first part of this important work Fuchs and Kramer examined time trends of physician market data over the period 1948 to 1966. The more startling results came from a cross-sectional study across states in 1966 using two-stage least squares estimation procedures. The simultaneous equation model is postulated on the basis of formulating intuitively plausible relationships, among observable variables, rather than deriving the relationships from the fundamental postulates of standard microtheory. It is not at all clear how the hypothesized model relates to one derived from utility and profit-maximizing behavior subject to constraints.

The four hypothesized behavioral equations are:

(i) per capita quantity as a function of price, per capita insurance benefits, median income, physician/population ratio, and the per capita number of hospital beds;

(ii) physician/population ratio as a function of price, output per physician, number of medical schools, number of beds, and per capita income;

(iii) output per physician as a function of price, number of doctors, and beds;

(iv) per capita insurance benefits as function of per capita output, price, a measure of immunization, income, and a ratio of health insurance premiums to benefits.

Such a model is very hard to interpret economically, even before a consideration of the estimation problems. Clearly, some variables are included in order to summarize the net effects of other, but related markets, for example, the ratio of insurance premiums to benefits, while other variables are clearly proxies, such as beds per capita, medical schools per capita, and so on. Finally, as the equations stand, they are ambiguous; for example, it is not clear that the first equation represents a demand equation or some type of partially reduced form equation.

For the purposes of this paper, the most important empirical result was that the physician/population ratio had a negative partial correlation with individual physician output rates, which stimulated the notion of physician-induced demand. Some very informal statistical analysis using British Columbia cross-sectional data by Evans [9], and Evans, Parish, and Sully [8] claims to be supportive of these results.

Some more recent cross-sectional work by Steinwald and Sloan [40] and Sloan [36] throws light on the sign of the partial correlation between the population/physician ratio and physician fees. In numerous other studies e.g., Newhouse [21], negative correlations were obtained, a result inconsistent with simple formulations of the standard model. The Steinwald-Sloan [40] study used micro data on physicians and defined population/physician ratios by county for general practitioners, but by state for internal medicine, pediatrics, and obstetrics-gynecology. For general practitioners and for general surgeons an increase in the ratio raises fees, as predicted by the standard model; but for the state-wide ratios for the other specialties, the opposite is the case. In Sloan [36], state data are used and mixed results are obtained.

One of the major deficiencies of the studies discussed above is the inadequate attention paid to the interaction of the insurance market with the market for physicians' services, a notable early exception being Fuchs and Kramer [16]. In a recent series of papers by Steinwald and Sloan [33, 39, 40] and a purely theoretical paper by Nordquist and Wu [24], the important, yet complex, role of the insurance market has been brought to light. Thus, while it is well known that insurance affects physician price-setting behavior, the effect varies considerably with specialty; the quantitative significance of insurance also varies with specialty, and insurance significantly changes the relative (patient perceived) prices of various procedures.

A recent book by Paul Feldstein [14] is devoted to the analysis of the supply of and demand for dental services. This work has been influenced substantially by the existing physician literature and broadly reflects these findings.

The Empirical Evidence of Production Rates, Quality, and Location

In an earlier paper Reinhardt [30] used a production function in which log output is a linear function of input levels and of log input levels. The functional form was chosen for its simplicity of estimation and flexibility in being able to fit a variety of economic hypotheses about productive relationships. One of the major empirical implications appears to be that complementary inputs to the physician in the physician firms are underutilized.

Scheffler [31] used the same functional form to estimate a dental production function. The results appear to be consistent with standard theory.

On the issue of the mobility of physicians and dentists, it would appear that, despite professional society spon-

sored barriers to entry at the state level, physicians and dentists relocate according to the predictions of standard theory (see for example, Benham, Maurizi, and Reder [3] and Guzick and Jahiel [17]). These results are consistent with the general findings of Pashigian [26] on the interstate mobility of all professionals.

The final relevant empirical issue discussed in the literature has been concerned with variations in the quality of visit insofar as this is measured by length of visit, length of waiting time, and length of traveling time. The last factor is potentially relevant as a partial explanation for the "perverse" signs between physician fees and the population/physician ratio. In any event examination of each of these issues shifts attention away from nominal physician fees towards real costs to the patient.

The major findings are that length of visit is positively correlated with physician/population ratios and that waiting time is negatively correlated (see Sloan and Lorant [35, 37]. A final interesting fact is that traveling time to doctors' offices in large urban areas appears to be longer than in smaller urban or suburban areas, although no allowance was made for differences in the distribution of specialties over regions [36]. However, the empirical evidence also shows that non-monetary factors such as travel time function as "prices" in discouraging demand, especially where the patient's nominal cost is low [1].

Summary

The most important empirical results for the purposes of this paper are that over time quantity of physician services is positively related to price; that a common, but by no means universal, result is that physician/population ratios are positively correlated with physician prices, but that waiting time and length of visit are negatively and positively correlated, respectively. Physicians appear to underutilize complementary inputs to the physician in the physician firm and an increase in the physician/population ratio decreases physician output rates.

PART II: THE THEORETICAL BASIS FOR AN EVALUATION OF SUPPLY-INDUCED DEMAND AND TARGET INCOME HYPOTHESES

The major objective of this section is to develop the theoretical framework within which the "supply-induced demand" and the "target income" hypotheses can be evaluated. Most generally stated, the problem of "testing theories" is a choice between competing theories. Consequently, the objective of this section can be restated as the evaluation of both the theoretical differences between the hypotheses and the implied empirical differences.

The first step in this task is to summarize the major criteria required for choosing between the competing hypotheses. The second step is the development of the appropriate formulation of what might be called the "standard neoclassical" model, that is, a model appropriate for the empirical analysis of physician and dental supply and demand functions. The final step involves a careful reformulation of the competing hypotheses in order to facilitate the choice between the hypotheses on both theoretical and empirical grounds.

Criteria for Choice between Competing Hypotheses

The criteria for choice between competing hypotheses involve both logical and empirical arguments. The logical criteria are concerned with questions of the logical consistency of each of the models and with questions on the scope of the empirical implications of each.

A preferred model is one that is internally consistent. By this is meant a model in which the various statements and sub-hypotheses of the model are not logically inconsistent with each other. As will be shown later in this paper, some variants of the competing hypotheses are internally inconsistent. Thus, it is useless to try to compare such a model with the standard model, which is internally consistent, until the former has been reformulated to remove such logical errors.

The second logical criterion concerns external consistency. New hypotheses are inevitably formulated, at least implicitly, within the context of a general theoretical structure. More precisely, new hypotheses are usually marginal modifications of an existing theory. The unchanged portions of the original theory are retained by the proponent of the competing hypothesis as part of the new theory. External consistency is the requirement that the formulation of the new hypotheses be logically consistent with those aspects of the old theory which are retained. As will be shown below, certain versions of the target income hypothesis are externally inconsistent. Again, comparing the empirical implications of alternative theories is of little use until the logical structures of both have been reconciled with the retained portions of the theory.

The two criteria just mentioned are absolute requirements that each competing hypothesis should satisfy before one would consider evaluating empirical differences. The next pair of criteria are relative in that one would use them as one aspect in the choice between the competing hypotheses. A preferred hypothesis is a more general hypothesis in that it relates to a wider variety of relationships between observable events than another hypothesis. The supply-induced demand hypothesis purports to achieve exactly that result.

The last theoretical criterion to be discussed is that the preferred theory makes more precise, or more refutable, statements. That is, an hypothesis which leads to only vague implications which would be consistent with a wide variety of different data would not, on logical grounds, be a preferred hypothesis to one that made precise statements that are consistent with only a narrow

range of alternative data. A simple illustrative example of a series of increasingly more precise, more refutable, and therefore theoretically more interesting statements is: quantity demanded and price are related, quantity demanded decreases (increases) when price rises (decreases), log quantity demanded increases proportionately to a decrease in price.

The second group of criteria involve the relationship between the empirical outcomes and the alternative predictions under the competing hypotheses. A first and necessary precondition to a useful statistical test to discriminate between hypotheses is that the models as formulated relate to the observed phenomena as specified by the theory. For example, using static competitive based hypotheses in a situation recognized to be one characterized by monopoly and dynamic characteristics vitiates any proposed test to discriminate between the competing hypotheses. Sometimes, the inapplicability is with respect to only one of the two hypotheses; nevertheless, any test to discriminate under such circumstances is still invalidated. As will be shown below, one of the problems involved in discriminating between the "target income" and "standard hypotheses" has been that with respect to the latter, very simple and obviously inappropriate versions have been used.

The most direct and often theoretically satisfying way of choosing between two hypotheses is to concentrate on those empirical implications under the two hypotheses which are clearly distinct; for example, under one hypotheses a given empirical situation yields a price increase while under the identical conditions the alternative hypothesis predicts a price decrease. Such tests might be termed crucial tests, since once one has verified that the hypotheses are relevant to the observed situation, the choice between the hypotheses is obvious on the basis of the outcome of any appropriate test. Unfortunately, as will be shown below, there are no crucial tests to discriminate between the competing hypotheses considered in this paper.

A less stringent requirement for choosing one hypothesis over another is that the preferred hypothesis not be rejected by some tests which reject the alternative and that otherwise both hypotheses withstand testing to an equal extent. For example, if hypothesis A meets some tests, but not others, and B meets some tests, but not others, and not necessarily the same tests in each case as for A, then one cannot on the basis of such tests alone choose between the hypotheses. But, if hypothesis A is not rejected by all tests not rejecting B, and further, A is not rejected by some further tests which do reject B, then at the appropriate confidence level one would be able to choose A over B. For a further discussion on these topics see [27, 28].

An Appropriate Version of the Standard Model for Analyzing Physician and Dental Supply and Demand

There are certain facts about the physician/dental market concerning which proponents of each competing hypothesis are in agreement. Physicians and dentists are self-employed entrepreneurs, either in terms of a single owner firm or in terms of a limited partnership. Thus, one must consider that the prime objective of each doctor or dentist involved in such an enterprise is to maximize his own preference function. However, insofar as market constraints limit the exercise of individual preferences, the maximization of profit for the doctor's firm might prove to be a relevant and useful sub-hypothesis. This is particularly the case where the physician and, to a lesser extent, the dentist face alternative employment opportunities as a physician or dentist.

Next, it is generally agreed that consumers are "ignorant" of the quality and physical nature of a doctor's services. What is at dispute is whether this ignorance is worse by some order of magnitude for the physician market than for almost any other market where demanders are non-specialists in the commodity being traded. Clearly, the effects of consumer ignorance are reinforced by the lack of advertising which, until recently at least, characterized physician and dental markets. However, to accept the intuitively plausible implication that such consumer ignorance leads inevitably to a situation in which each doctor is able to extract extra rent from that ignorance is to engage in a *non sequitur*.

Without at this time going into details, we might note several factors that modify the naive conclusion. If consumers are ignorant, risk aversion will lead them to engage in costly activities to alleviate that ignorance by checking medical opinions with other doctors, by reading medical books and popularized versions of medical books, by exercising caution in taking action, for example by delaying an operation until the patient is convinced it is necessary, by reserving the right to sue for physician errors, and so on. Secondly, if there is consumer ignorance and lack of advertising inhibits consumer arbitration of physician fees, then physicians themselves are ignorant of the demand curve, which they would otherwise face, especially if the market is assumed to be monopolistically competitive. Physician market ignorance and risk aversion may lead to lower prices being charged than would maximize expected profit. The immediate result of the assumption of ignorance is that the market will be characterized by a much greater dispersion of real (quality compensated) prices for any given set of market conditions so that discrimination between the hypotheses is made more difficult. Without further and more complete analysis one cannot predict whether the expected physician fee will be higher or lower than the expected fee without ignorance.

Another commonly accepted fact is that because of real transportation costs the relevant physician/dental markets are contained within very small geographical regions for all but the most specialized services. Indeed,

it is generally agreed that the geographic extent of the market increases with the degree of specialization and decreases with population density. Consequently, one must be careful in cross-sectional analysis to avoid problems introduced by specialists having geographically broader markets than general practitioners and to assure that regional differences have been arbitraged away as they would be if transportation costs and barriers to factor mobility were negligible. An obvious, but by no means insignificant example, is the difference in office rentals faced by doctors in different sized communities.

While the above "facts" are widely regarded as being valid, a number of other aspects of physician supply have been generally ignored by the proponents of each hypotheses. The most striking and perhaps the most difficult to handle is the effect of technological change, not just on the productive process, but also on the range of options in demand. The former aspect is relatively straightforward and has received some attention in the literature. The latter aspect has been almost totally ignored. Yet in any time-series analysis over any period of more than a few years, especially over the past twenty, the increase in the number, variety, and quality of services available to the consuming patient has been substantial. In addition, the development of drugs, antibiotics, inoculation and vaccination procedures, has dramatically changed the nature of the basket of health services demanded from the physician and dentist. In any empirical analysis using historical data one must be careful to allow for the effects of such changes.

This factor has cross-sectional implications as well in that for various easily recognized economic reasons the geographic distribution of doctors is not homogeneous with respect to specialty. Consequently, in cross-sectional analyses across geographical areas that involved aggregation within such areas for price indices and so on, one must be careful to check that the distribution of specialties within each geographic region is the same across regions. For example, a simple comparison of urban/rural physician fees must allow for the large difference in the percentage of specialties between urban and rural areas.

Finally, on the demand side, while almost all researchers have recognized that demand behavior with insurance coverage differs from that which applies without coverage, few have proceeded to incorporate the cost of insurance in the consumer's decision and to allow for the consequent endogeneity of the proportion of the population that has insurance. Further, the role of insurance in the physician market is more complicated than has been generally allowed in the analyses to date. The two difficulties are the differential effects of coverage across specialties, general practitioners being the least affected in general; and the effect of the implicit change in relative prices between those physician services covered and those not covered on the distribution of services in the representative basket used to calculate price indices.

The following comments provide a set of empirical circumstances that will indicate the ostensibly appropriate model to be used and they indicate a number of warnings to be kept in mind when the models are to be tested.

Consequently, as a first approximation, let us consider a comparative static equilibrium model of the owner/entrepreneur (or doctor) firm, wherein each doctor or dentist has the opportunity to be employed as a physician or dentist at a market determined wage rate. Let us also assume as a first approximation that factor inputs to doctor/dentist firms are competitively determined, that market ignorance and hence uncertainty and risk are non-existent, and that technology is constant. The model described below is a modest modification of one developed by Olsen [25]. The unit of analysis is the individual physician or dentist. As usual the model is formulated with a view towards facilitating the analysis of the major issues and suppressing details that, while of importance in and of themselves, are not central to the current argument. However, once one considers the estimation of parameters and the testing of hypotheses, the numerous simplifying assumptions must be re-examined.

The individual physician is assumed to have a preference ordering over commodities, $y_1, \ldots y_m$, and leisure L, represented in the usual manner by:

$$U(y_1, \ldots y_m, L). \tag{1}$$

On the assumption that the relative prices of goods remain constant and that each physician consumes an insignificant amount of his own product, y_1, equation (1) can be rewritten conditional on final goods prices as:

$$U(M, L), \tag{2}$$

where M represents income.

The production function for medical services can be formulated as:

$$y_1 = f(x_1, \ldots x_n), \tag{3}$$

where $x_1, \ldots x_{n-1}$ represent quantities of inputs to the productive process of a doctor's provisions of services and x_n is the level of doctor inputs measured in hours of effort at some given average intensity. In this first approximation to the doctor firm, physician input by one doctor is assumed to be perfectly substitutable for physician input by any other doctor. Let w_n represent the market determined opportunity cost for the self-employed physician, and w_i, $i = 1,2, \ldots n-1$ are the market prices for all other inputs. Let $T = L + x_n$ denote the doctor's time constraint.

The doctor's income from operating his own firm is given by:

$$M = \pi + w_n x_n \tag{4}$$

where

$$\pi = p_1 y_1 - \Sigma_1^n w_i x_i \tag{4'}$$

is the firm's profit and the doctor's wage income from his own firm is $w_n x_n$.

For each level of doctor input, x_n, we can consider the cost-minimizing demand for inputs and the cost-minimizing supply of output as a function of x_n. Thus:

$$\begin{aligned} x_i &= g^i(x_n \langle w_i \rangle),\ i = 1,2 \ldots n \\ y_i &= f(x_n \langle w_i \rangle),\ \langle w_i \rangle = w_1, w_2, \ldots w_{n-1} \end{aligned} \tag{5}$$

and the profit function can now be written as:

$$\pi(x_n, p_i) = p_i f(x_n) - \Sigma_1^{n-1} w_i g^i(x_n) - w_n x_n, \tag{6}$$

where we assume for now that services produced according to equation (6) are in fact consumed by patients at the price p_i. Equation (6) can be summarized by:

$$\pi(x_n, p_i) = p_i f(x_n) - C(x_n) - w_n x_n. \tag{6'}$$

So far the individual doctor's preference function has played no role in the analysis. There are two polar alternatives at this stage: either there is an efficient market for doctor's own firm with easily varied hours, or doctors are restricted to providing physician services only through single doctor firms. In the analysis in this paper, the former will be assumed as the more realistic and relevant assumption about physician behavior. If we also assume that there are no further unconsidered benefits or costs to being self-employed, then the utility-maximizing behavior of each physician is to operate the firm at the profit-maximizing input level of physician input. There are two cases: either hire the extra physician input to the firm if the doctor/owner wishes to work less than the optimal (profit-maximizing) number of hours, or the doctor/owner hires himself out for extra income if he wishes to work more than the optimal number of firm hours. Thus, the "efficient doctor market" assumption leads to a dichotomy between the profit-maximizing output level of doctor firms and the individual utility-maximizing hours of work.

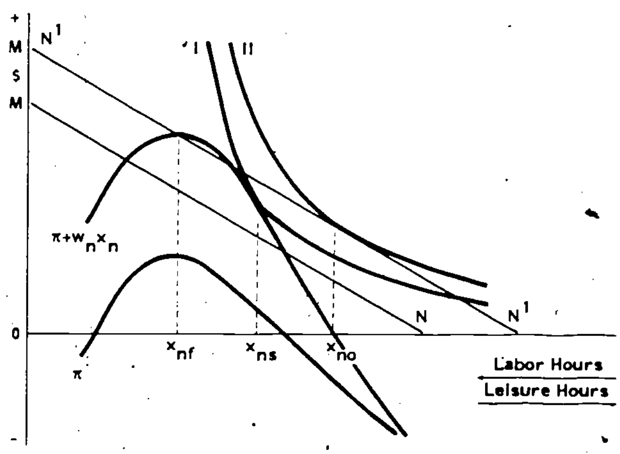


Figure 1.
Optimal Firm and Individual Physician Hours

The short-run optima can be easily plotted as shown in Figure 1, provided the optimum size of a doctor firm is not more than T hours. [Even if the optimum number of doctor hours is greater than T, the general result still holds, but a more complex diagram is needed to illustrate the argument.]

The line MN represents the income/time constraint line, i.e., $M = w_n T$ and the point N represents $T = L$. The lines I, II represent indifference curves. π is the firm's short-run profit function expressed as a function of x_n physician hours of input; x_{nf} represents the optimal (profit-maximizing) number of hours of physician input. x_{n0} represents the doctor's personal utility-maximizing hours of work with leisure given by $L = T - x_{n0}$. The point of maximum utility is reached by the doctor operating the firm at x_{nf}, working himself in the firm for x_{n0} hours, and hiring a colleague (or taking a partner) to work $(x_{nf}-x_{n0})$ hours. If doctors were restricted to providing service through single doctor firms, the utility-maximizing position would be at x_{ns} and the firm would not maximize profits in that the owner/doctor would have to trade profits for leisure.

The analysis so far has ignored the issue of whether the doctor firm is monopolistically competitive or the doctor is a price taker. While this distinction is important for some considerations, it is of little moment in the current discussion. In long-run equilibrium in either case, the individual small firm's optimum profit level is forced to zero and the highest attainable time/income constraint line reduced to MN.

The above analysis also illustrates a more general comment. While an individual doctor's preference function is certainly more complex than the simple assumption that income and leisure are the only two inputs to it, the efficient physician market assumption implies that doctor firms (as opposed to doctors) will be operated efficiently, even if individual doctors wish to engage in philanthropy, pursue interesting cases, etc. Thus, the extent to which the efficient physician market hypothesis is violated will determine the extent to which doctors will be able and willing to trade profits for the consumption of philanthropy, scientific interest, hobbies, and whatever nonprofit generating pursuits researchers may consider physicians desire. The efficient market hypothesis does not deny that physicians will engage in such pursuits, but states that the pursuit will be separated from the issue of maximizing firm profits. Market ignorance will further modify these results through the introduction of risk and uncertainty on both sides of the market. Such refinements are not usefully discussed in the current context.

While the above model of physician/firm behavior is far from being a realistic model, that is, it is not suitable for testing as it stands, it does provide the appropriate standard theoretical formulation with which the models of the new hypotheses can be usefully compared.

Little needs to be said about the relevant demand functions for medical and dental services. Both sets of proponents agree that the standard theory applies with the exception of the modification introduced by the new hypotheses which are to be discussed in the next section. However, there seems to have been some confusion over two issues.

First, as mentioned above, the number (or proportion) of the population insured is an endogenous

variable in any market-wide analysis, or analysis using broad aggregates over time. As is readily understood, the decision to buy medical insurance, either individually or jointly through an employee group, is the outcome of weighing the certain costs of insurance against the expected value of the benefits. Risk-averse behavior by consumers enables firms to sell insurance at a price that covers costs of provision as well as benefits. Consequently, even if medical insurance were completely monopolized, the profit-maximizing insurance firm has an incentive to keep the price of medical services down, albeit perhaps not to the extent that a competitive insurance market would. The reason is that through the insurance firm the consumer is buying the product "maintenance of good health." An insurance company, by holding down medical costs, holds down the cost of the most important element in its health maintenance product. In any case, analyses of physician markets must also consider the interaction of that market with the market for medical insurance.

The second issue over which there has been some confusion concerns the relationship, if any, between size of population and number of physicians. If a comparison is to be made, one can only do so in the context of equilibrium, within which one can consider the change in the equilibrium number of physicians to an increase in population, shift in the patient demand function, or a shift in physician production costs. In the simplest competitive situation, equilibrium is defined by:

$$y_1^* N_D = q(\text{Pop}), \qquad (7)$$

where $y_1^* = F(x_n^*)$ is the optimum size of doctor firm working at the optimal physician input rate of x_n^*, N_D is number of physicians, $q(P_1)$ is the mean patient demand function, and Pop is the size of population. The formulation of the equilibrium equation above assumes that $q(P_1)$ is invariant to shifts in Pop. The main problem in comparing the equilibrium solutions of N_D for given Pop is to take appropriate account of the shifts in both $q(P_1)$ and y_1^*. Only in the simplest of circumstances would one expect the ratio of $N_D/(\text{Pop}) \equiv q(P_1)/y_1^*$ to remain invariant to changes in Pop; in any case, the question can only be answered even in the competitive case after one has evaluated the long-run effect of changes in demand on y_1^*, the optimal output rate. More likely, $q(P_1)$ is itself a function of Pop.

Clearly, N_D is an endogenous variable determined by the interaction of market actions by both demanders and suppliers and the long-run effect of entry and exit. This somewhat obvious statement is not nullified if professional associations, such as the A.M.A., are successful at limiting entry of physicians to any given geographic region. The effect of such restrictions is to raise the opportunity cost to the self-employed physician by raising the rent paid, as reflected in the increase in w_n, to a specific input to the physician firm, namely physicians. In short, restricted entry raises w_n, tilts the curve MN in Figure 1 up, shifts the cost curves $g'(\cdot)$ and the function $y_1 = f(x_n)$. Under standard assumptions about the productive process, the profit function π will be shifted down and to the right so that the optimum output rate, y^*, will fall, but the optimum physician input level x_n^* will fall proportionately more. Consider for example Figure 2 which illustrates this brief argument.

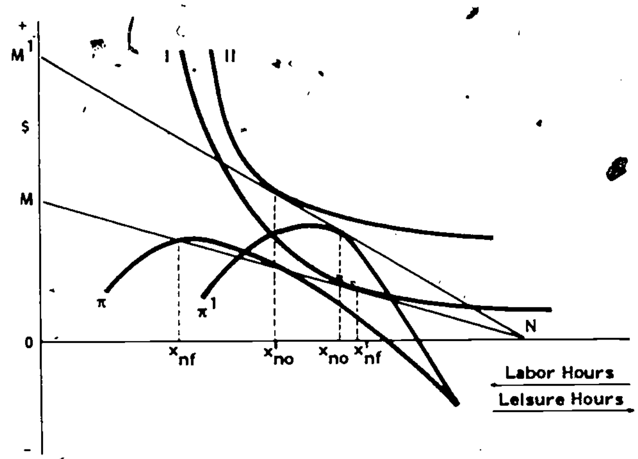


Figure 2.
Potential Effects of Physician Restricted Entry

Suppose an increase in restriction of entry leads to a shift from the long-run equilibrium position given by x_{nf} and x_{no}, the optimum firm input level and the optimum working rate at w_n to the new long-run equilibrium optima at x_{nf}^1 and x_{no}^1 at the rent increased wage rate of w_n^1, $w_n^1 > w_n$. [The new time constraint line is MN^1]. In this example, we observe that the effect has been to decrease the optimal physician hours of input to the firm, but to increase the optimum number of hours of work by each physician.

Similar comments apply if physician firms are monopolistically competitive; the main difference between the two cases is in the definition of the optimum output level, y_1^*, which will be less than the corresponding competitive optimum rate. Equation (7) still applies as does Figure 2. However, the evaluation of circumstances under which the equilibrium ratio of N_D/Pop $= q(P_1)/y_1^*$, where $P_1 = P_1(q)$ remains constant, becomes more algebraically complex.

An important aspect of the current formulation of the standard theory in this situation is that restricted entry leads to an increase in physician or dental rents, but does not lead to any concept of excess demand for physician or dental services. This topic in particular will be reintroduced and elaborated in section III.

The standard theoretical framework is now in a form suitable for evaluating the different implications of the new competing hypotheses to which attention will be directed. There are two main alternative hypotheses: the "physician ability to influence demand," which is labeled Evans Model I (after one of the earlier proponents), and the so-called "target income hypothesis," which is labeled Evans Model II. The first task is to reformulate both models in a manner that cap-

tures the major ideas in the new hypotheses without suffering from the logical deficiencies mentioned above.

A Theoretical Analysis of Evans Model I (Physician-Induced Demand)

Evans model I, most formally introduced in [9], and further examined with some attention to formal analysis by Sloan and Feldman [38], is based on an intuitively appealing idea: consumers are ignorant of their medical "requirements" and can be persuaded to pay more for more services whenever the physician finds it in his interest so to persuade the patient. Over time, the earlier more forceful statements (and therefore more empirically meaningful) have been modified considerably. What might be regarded as a current position is summarized by Evans [7, p. 21]:

> In influencing demand it is not necessary for physicians to turn away critically ill patients or to hoodwink the healthy in defiance of professional ethics; all that is needed is that more time, effort, and care be spent with each presenting patient when the apparent exogenous workload is reduced.

Depending on how this statement is interpreted it comes dangerously close to saying nothing at all. For example, if fees are paid per visit, the physician maintains certain set office hours, and the physician's "extra care" does not result in extra visits, then the phenomenon, if it exists, has little effect on observable events, unless one observes length of visit. Even so, this model is still very close to a standard model under similar circumstances in that one way to vary real prices of service at lower transactions costs than by varying nominal prices directly is to vary the quality of the product, especially in response to temporary variations in demand.

There is the further difficult problem of identification of the Evans effect in distinction from the physician's normal provision of information about the product ("normal," in that a major part of any salesman's role is to provide information about his product). In physician firms, doctors are usually their own salesmen.

In the subsequent theoretical discussion, this identification problem will be ignored. However, if empirically a significant effect of "demand shifting" is observed, one must worry about separating the Evans effect from the supply of information effect.

Evans model I as presented by Sloan and Feldman [38, p. 7] is:

$$
\begin{aligned}
&U(Y,W,D)\\
&W = R\cdot f(P,D) \qquad\qquad (8)\\
&Y = R\cdot f(P,D)\cdot P - C(W),
\end{aligned}
$$

where $U(\cdot)$ is a preference function for the individual physician, Y is income, W is "workload," D is physician's discretionary influence on patient demand, R is the population-physician ratio (assumed to be exogenous), P is physician fee, $f(\cdot)$ is the patient demand function, and $C(\cdot)$ is the cost of providing services. As Sloan and Feldman note [38, p. 8], there is no clear statement as to what "D" represents or how it would be measured; consequently, unless arbitrary assumptions are made about the partial derivatives of $U(\cdot)$ and $f(\cdot)$ with respect to D and the cross-partials involving D, no theoretical conclusions can be drawn.

While W is called "workload," it is really a measure of quantity demanded per physician. The implicit assumption in this form of Evans model is that all demands at given P and D are met by the physician firm. This assumption notwithstanding, W is still an inappropriate argument in the preference function. One of the crucial missing links in this formulation of the hypothesis is some form of production function which shows how physician hours of work are related to output. Through a time constraint one could then introduce hours of work into the preference function by using the equation $L = T - x_n$, where L is hours of leisure, a variable that does occur in the preference function, and where x_n is hours of work.

With these difficulties in mind, consider the following modification of the standard model so as to incorporate the Evans model I effect.

Let us define D in terms of the hours of effort expended in persuading patients they need more health care; D can be observed and measured, at least in principle. The demand curve faced by the physician firm can be written:

$$p_1 = p_1(f(x_n),D), \qquad (9)$$

where we are assuming that the physician firm is in a monopolistically competitive market, that the quantity demanded at price p_1 given D is in fact supplied (this assumption justifies using $f(x_n)$, which is firm output in the demand equation), and that an increase in physician efforts to shift demand is successful, but only at a a decreasing rate; i.e., $\partial p_1/\partial D > 0$, $\partial^2 p_1/\partial D^2 < 0$.

On the physician supply side, let us continue the useful simplifying assumption that there is an efficient physician market, so that in terms of the physician firm's output, one need only consider profit-maximizing behavior. Let x_{n1} denote physician hours in firm production and x_{n0} the physician's personal (own) time spent working. The profit function is now:

$$
\begin{aligned}
\pi = p_1(f(x_{n1}), D)f(x_{n1}) - C(x_{n1})&\\
- w_n D + x_{n1}& \qquad (10)
\end{aligned}
$$

and the physician's income is given by:

$$M = \pi(x_{n1}) + w_n x_{n0} \qquad (11)$$

with time constraint: $(T-D) = x_{n0} + L$. Compare equations (10) and (11) with equations (4) and (6'). Under the efficient physician market hypothesis, the doctor determines the firm's profit-maximizing physician input level and his own utility-maximizing workload separately; the difference in optimal physician hours is cleared in the open physician market as discussed above. The profit function is maximized with respect to x_{n1} and D and the utility function with respect to M and L, where M is a function of x_{n0} given $\pi(x_{n1})$ and

D. For each given value of D the analysis is the same as that shown in Figure 1. For notational convenience, the subscript "f" on x_{nf} can be dropped without confusion, since the ensuing analysis need deal only with finding the optimal value of π.

Letting $H_1(x_n,D)$, $H_2(x_n,D)$ denote the normal equations obtained from the first order conditions, one has:

$$\frac{\partial \pi}{\partial x_n} = \frac{\partial p_1}{\partial f}(\cdot)\frac{\partial f}{\partial x_n} f(x_n) + p_1(\cdot)\frac{\partial f}{\partial x_n} - \frac{\partial C}{\partial x_n}(\cdot) - w_n = 0$$

or $$H_1(x_n,D) = MR(x_n,D) - MC(x_n) = 0;$$

$$\frac{\partial \pi}{\partial D} = \frac{\partial p_1}{\partial D}(\cdot)\, f(x_n) - w_n = 0, \tag{12}$$

or $$H_2(x_n,D) = p_{1D} f(x_n) - w_n = 0.$$

$MR(x_n,D)$ and $MC(x_n)$ denote rates of increase in total revenue and costs, respectively, with respect to changes in physician input. In line with the current literature in this area let us assume that the above normal equations define a regular maximum so that the usual second order conditions hold.

If we now wish to consider the effects on this model of shifts in certain parameters of the system, we can do this by examining the signs of the coefficients in the following equation:

$$\frac{dx_n}{dD} = (-1)\begin{pmatrix} H_{1x} & H_{1D} \\ -H_{2x} & H_{2D} \end{pmatrix}^{-1} \begin{pmatrix} H_1\lambda \\ H_2\lambda \end{pmatrix} d\lambda, \tag{13}$$

where $d\lambda$ represents the increase in some parameter λ, $H_1\lambda$, $H_2\lambda$ represent the partials of $H_1(\bullet)$ and $H_2(\bullet)$ with respect to λ, and the remaining symbols have the usual interpretation. Equation (13) can be usefully rewritten as:

$$\begin{pmatrix} dx_n \\ dD \end{pmatrix} = (-1)\mathrm{Det}^{-1} \begin{pmatrix} H_{2D} & -H_{1D} \\ -H_{2x} & H_{1x} \end{pmatrix} \begin{pmatrix} H_1\lambda \\ H_2\lambda \end{pmatrix} d\lambda, \tag{13'}$$

where:

$$H_{1x} = \frac{\partial MR(x_n)}{\partial x_n} - \frac{\partial MC(x_n)}{\partial x_n} < 0$$

$$H_{1D} = \frac{\partial MR(x_n)}{\partial D} > 0 \tag{14}$$

$$H_{2x} = H_{1D}$$

$$H_{2D} = P_{1DD} f(x_n) < 0$$

The signs of H_{1x} and H_{2D}, and that Det (the determinant of the matrix of partials) > 0 are a consequence of the assumption that the profit function has a regular maximum. That H_{1D} (and by symmetry H_{2x}) is positive in a neighborhood of the maximum is an assumption, albeit a plausible one. The assumption is useful because it enables one to derive unambiguous conclusions about signs of certain changes and because it is favorable to the Evans model.

The above formulation enables one to compare the theoretical predictions of the Evans model to those of the corresponding standard model; in the current formulation the standard model results are obtained by setting all partial derivatives with respect to D to zero.

Consider an increase in w_n, the firm's opportunity cost for physician input. Since both H_{1w}, $H_{2w} = (-1)$, one immediately concludes that both dx_n/dw and dD/dw are less than zero, certainly an intuitively pleasing result. Alternatively, consider a shift in the production function $f(x_n)$. Under these circumstances $H_1\lambda < 0$ and $H_2\lambda = 0$, so that both $dx_n/d\lambda$ and $dD/d\lambda < 0$. Suppose now that an increase in λ represents a shift in demand due to a population increase, or an increase in income, or whatever. Under this circumstance, $H_1\lambda = \frac{\partial MR(x_n,D)}{\partial \lambda}$ which is clearly positive under the assumed conditions. $H_2\lambda = \partial P_{1D}/\partial\lambda > 0$ is a correspondingly plausible assumption that strengthens the case for the Evans model. Under these assumptions, one concludes once again that both $\partial x_n/\partial\lambda$ and $\partial D/\partial\lambda > 0$.

Thus, as these examples illustrate and as we can conclude from a close inspection of equations (12), (13), and (14), parameter shifts that yield an unambiguous change of sign in x_n produce the same change of sign in D. More importantly, at this level of analysis there are no qualitative differences in observable behavior between the Evans model and the standard model.

Let us now consider the effect of parameter shifts on the change in equilibrium price, a matter of some importance in these models. Formally, we may write:

$$\partial P_1/\partial\lambda = P_{11}\frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial \lambda} + P_{1D}\frac{\partial D}{\partial \lambda} \tag{15}$$

The sign of $(P_{11}\frac{\partial f}{\partial x_n})$ is negative and that of P_{1D} positive, so that an unambiguous effect on P by some shift λ can be determined in this model only if $\partial x_n/\partial\lambda$ and $\partial D/\partial\lambda$ are of opposite signs or one of $\partial x_n/\partial\lambda, \partial D/\partial\lambda = 0$; but the previous analysis concluded that $\partial x_n/\partial\lambda$ and $\partial D/\partial\lambda$ are of the same sign in general. In the standard model $\partial P/\partial\lambda$ is of opposite sign to $\partial x_n/\partial\lambda$. Thus, in this situation the standard model provides a refutable statement, but the Evans model does not. Thus, if the predicted sign of $P/\partial\lambda$ were found to be inconsistent with a set of data, then that result provides evidence against the standard model, but provides no evidence, either for or against, the Evans model. Consequently, one cannot use such a test to try to discriminate between the two hypotheses.

The difficulty in all this is, of course, that as shown in equation (15), for any change $d\lambda$, the productive effort effect $(P_{11}\frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial \lambda})$ is offset by the demand shift effect $(P_{1D}\frac{\partial D}{\partial \lambda})$.

An examination of Figure 3 indicates one class of methods whereby potential discriminatory tests can be constructed. Let us assume $d\lambda$ represents a shift down in

the marginal cost curve which leaves $f(x_n)$ invariant. Recall from the above discussion that $\partial D/\partial\lambda < 0$.

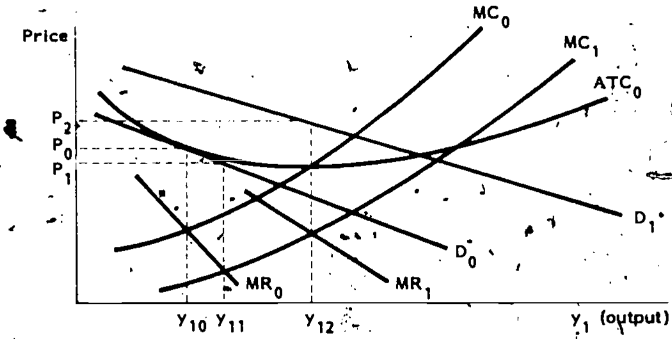


Figure 3.
A Comparison of Equilibria under Competing Hypotheses

Suppose that at some point in time, t_0, the physician firm market is in long-run equilibrium so that each firm's equilibrium price and output are as shown in Figure 3, P_0, y_{10}. Now suppose that there is a decrease in marginal cost from MC_0 to MC_1. Under the standard model the predicted short-run equilibrium response is to produce at y_{11} at a price of p_1. This standard response can be predicted for each firm, even if the Evans effect is present, under the following assumptions:

(i) the slope of the ATC curve is known at y_{10}
(ii) the functional form of the demand curve D_0 is known and the demand function depends on at most two parameters (where D is expressed as a function of price only).

With the above information (the information required in (i) can often be obtained separately by cost function estimation), one can estimate the function $D_0(y)$ and hence derive the corresponding MR(y) curve. Knowing the shift in MC and the functional relationship between cost and output enables one to predict Y_{11} and P_1.

If the D effect is present, then actual output and price even under short-run equilibrium will be given by y_{12} and P_2 where $y_{12} > y_{11}$ and $P_2 > P_1$. The relevant curves are D_1 and MR_1. Thus, while one cannot predict that $P_2 > P_0$, one can predict $P_2 > P_1$. Similar tests can be constructed with respect to other types of parametric shifts. The most important criterion for developing a successful test is the ability to predict the short-run equilibrium result after the parameter shift, if the D effect is not present.

Figure 3 is also useful in illustrating the effects on the standard model of incorporating the Evans effect. With a decrease in marginal costs, output rises in both models and while price falls in the standard model, price may or may not fall with the Evans effect, the net result depending on the relative degree of shifting of the curves.

The cost function, or rather, the function relating cost to physician input rate, can be used to provide an indirect test of the Evans effect. From the standard analysis discussed above one can derive the function $C = C(X_{nf})$ and if one can obtain observations on C at various output levels for constant input prices, such that the cost function is identified, one can test indirectly for the presence of the Evans effect. Note that one of the problems with trying to devise tests on the effects of D is that in general D is unobserved; what is observed is $X_n = X_{nf} + D$ and $C(\cdot)$ is a function of X_{nf} alone. Further, from the analysis above it is clear that X_{nf} and D are positively related; that is, to any parameter shift the directions of change of X_{nf} and D are the same. Consequently, if the D effect is present in any regression, one must use X_n, not X_{nf} as the regressor. As a consequence, the regression model is misspecified in a fairly precise manner so that specification error tests can

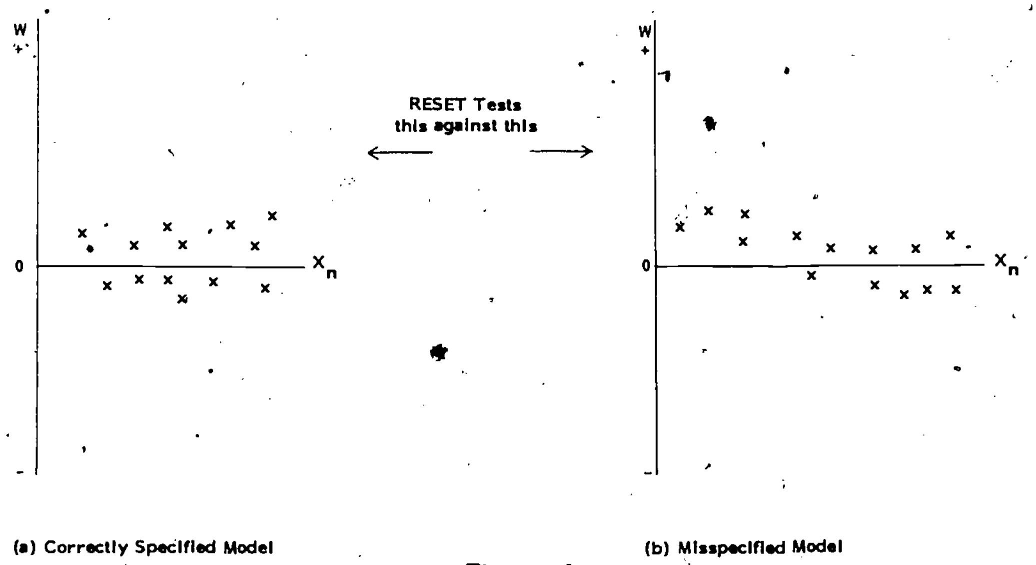


Figure 4.
Illustration of the RESET Specification Error Test.
A Comparison of Plots of Disturbances Terms Against X_n

be used to discriminate between the models; see, for example, Ramsey [27, 28, 29].

As an illustration of this approach, suppose that the theoretical cost function can be written:

$$C = a_0 + a_1 x_{nf} + a_2 z + u \qquad (16)$$

where x_{nf} is physician input, z represents other variables which might have to be included in the regression equation given the sampled observations, and u is the disturbance term. The actual regression run is:

$$\begin{aligned} C &= a_0 + a_1 x_n + a_2 z + w \\ w &= u - a_1 D \end{aligned} \qquad (17)$$

and w and x_n are correlated with correlation given by $-a_1(\mathrm{Var}(D) + \mathrm{Cor}(X_{nf}, D))$ so that the correlation coefficient has the opposite sign to that of a_1, which is positive. Similarly, the relationship between x and w can be determined from the above relationships and the easily verified relationship between x_n and z. These relationships enable one to determine the effect of the induced bias on the estimates of the coefficients. More importantly, one's first step is to test the model for specification error by means of the RESET test (Ramsey [27]).

Essentially, the RESET specification error test is an F type test on transformed residuals to test the hypothesis that the disturbance term vector is distributed as normal with null mean vector and scalar covariance matrix versus the alternative hypothesis that the disturbance term, while normally distributed, has a non-null mean. Figure 4 illustrates the idea. Under the specification error mentioned above, the expected value of the disturbance term is a decreasing function of x_n, hours of physician/dental input.

The comparative analysis of Evans model I can be easily summarized. First, far from the Evans model leading to startlingly different conclusions as was claimed in some earlier work, the difficulty in discriminating between the hypotheses is that there is so little qualitative difference. Although Evans model I, when properly reformulated, can be regarded as a generalization of the standard model, it is less precise in that it provides fewer refutable hypotheses. More importantly, the Evans model does not and indeed cannot explain the partial correlations which have been thought to be inconsistent with the standard model, most noticeably the positive correlation between price and physician/population ratios. The further difficulty of the non-observability of D leads one to conclude that in the absence of strong empirical evidence in favor of the Evans hypothesis the standard model is preferable to Evans model I.

In this connection, two tests have been proposed, which, if performed under the specified conditions, would enable one to discriminate between the standard model and Evans model I.

A Theoretical Analysis of Evans Model II (Target Income Hypothesis)

The most complete and formal specification of what might be called Evans model II is to be found in Evans [7]. One suspects that Evans model II was originally conceived to improve model I's disappointing predictive performance. The original version was the stronger is that target incomes of physicians were posited to be fixed over time, even if variable over physicians; see Evans [7]. This strong form also seems to have been quickly shelved, partly because there seemed to be no empirical evidence in favor of fixed money incomes. While there has been much loose discussion of Evans model II, the so-called "target income" hypothesis, there seems to have been no careful formal exposition of this idea. The most detailed and explicit account of this concept is contained in Evans [7]. The basic framework of that exposition is worth setting down, since the principal reference is somewhat obscure to non-Canadian readers.

There are four behavioral equations and a number of definitional equations. It is important for the reader to note that the behavioral relationships are postulated as presented below and are not explicitly derived from any idea of individual optimization subject to constraints.

The mean patient demand function is:

$$q = q(s, P, \mathrm{IPop}/\mathrm{Pop}, t) \qquad (18.1)$$

where s is the "workload" surplus defined by actual workload minus desired workload, P is price, IPop/Pop is the proportion of population with insurance, and t is time. Evans assumes $\partial q/\partial s < 0$, that is, if the actual workload exceeds desired workload, the more physicians will strive to lower quantity demanded.

Desired workload, W_d (actually desired output since workload W is defined by W = (qxPop)/MD, MD = number of physicians) is assumed to be some function of the physician fee (P) and the physician's income from the medical firm (N). The relationship between W_d, P, and N is supposedly meant to be derived "in the customary way predicted by the work-leisure trade-off," but Evans then assumes $\partial W_d/N < 0$ and $\partial W_d/\partial P > 0$ for any P and N; in short, one has:

$$W_d = W(N, P) \qquad (18.2)$$

There is a price equation:

$$P = P(N, N', S) \qquad (18.3)$$

where N' is the physician's target income which is assumed to vary *both over time and over physicians in a non-specified manner.* Evans assumes $\partial P/\partial S > 0$ and P is reduced if (N – N') increases. This is the only equation in which N' occurs.

The final behavioral equation is:

$$MD = M(N, S) \qquad (18.4)$$

where $\partial M/\partial N > 0$ and $\partial M/\partial S < 0$, MD is the number of physicians within the region under consideration. The idea is that "high" net incomes encourage immigration whereas "workloads" greater than desired discourage immigration.

The remaining equations are essentially definitional.

Regional quantity of medical services demanded, $Q = q\mathrm{x}\mathrm{Pop}$

Gross physician income, $G = PxW$

Net physician income (after allowance for costs), $N = (1\text{-}c)G$ (18.5)

Fraction of gross income going to associated inputs, $c = c(W,P)$

After re-expressing the above equations in log-linear form, Evans sets down a comparative statics simultaneous equation model of the type:

$$Hu_t = Fx_t \qquad (19)$$

where H, F are matrices of coefficients, $u_t' = (LnP_t, Lnq_t, LnN_t, LnS_t)$ is the vector of natural logarithms of the values of the endogenous variables and $x_t' = (LnN_t^I, Lnt, Ln(Pop/MD)_t, Ln(Pop/MD)_t)$, (*sic!*), are the exogenous variables. On the basis of the above model and on the further assumptions that the price elasticity of demand goes to zero (with increases in insurance coverage) and that the workload surplus variable, S, has no effect on price (!), Evans derives the expected sign changes one would get by solving for u_t in terms of x_t.

From Evan's perspective the most important result is to show that an increase in the ratio (Pop/MD) reduces P and q, but raises N. From this statement Evans concludes that a policy that increases the number of physicians and hence lowers the ratio (Pop/MD) will under these circumstances be counterproductive in that price and per capita expenditures will rise, even though physician incomes will fall (Evans [7, pp. 29, 30]).

Let us now examine this seemingly plausible model more carefully. For all its apparent sophistication the model is seriously flawed logically. Much of the difficulty, one suspects, stems from the author trying to write down a series of observable relationships with appealing partial derivatives without *deriving* his relationships from first principles. The best way of analyzing the logical difficulties of this model is to try to derive it from individual preference functions after due allowance for the modifications Evans wishes to introduce. However, one need not be so formal in order to begin to see the logical difficulties in this exposition.

The first difficulty occurs with the formulation of the demand curve. Besides the *ad hoc* and essentially careless way in which the role of insurance coverage is handled, S is postulated as an argument of the demand equation, but S is the surplus workload as *perceived by physicians (sic)* [7, pp. 22, 30]. Clearly, S itself cannot be an argument in q(•). One way out of this difficulty is to specify a further relationship: D = D(S), where D is the doctor's hours of efforts in persuading patients to buy more medical care. Presumably D(0) would then represent the equilibrium level of effort. However, given the analysis in the previous subsection, the relationship between D and S is not one that is postulated in a theoretical vacuum, but should be derived from the individual physician's efforts to maximize profit and utility as already discussed. The introduction of S in q(•) is only the first of several clues to the suspicion that Evans model II is not an equilibrium model, but a dynamic adjustment model, or more accurately a dynamic non-adjustment model.

W^d, labeled desired workload, is said to be derived in the usual way from the work/leisure trade off. First, as noted in the previous subsection, W and hence W^d are not workload measures, but measures of physician firm output, quite a different concept. W^d can be regarded as that output rate determined by $f(x_n^d)$, where $f(\cdot)$ is the production function and x_n^d is the profit (or utility) maximizing level of physician input. The previous analysis demonstrated how x_n^d could be derived. From these relationships one could then derive the equilibrium relationship between the optimal values for price and $y_t = f(x_n)$. This procedure, however, shows that W^d (y_t in the notation above) as a function of P and N (or M in the previous notation) is not a fundamental behavioral relationship, but one derived from the solution of the optimizing conditions. With this in mind, it is clear that $W^d = W(P, N)$ is an incomplete specification of the variables affecting the equilibrium level of output.

Furthermore, Evans has postulated that physician preference functions are such that for each physician utility is maximized when actual income is equated to "target income." But this implies in turn that W^d (or Y_t) is a function of N^I as well. Thus, one can only conclude that the derived relation $W^d = W(P, N)$ is also misspecified.

The introduction of a target income N^I which varies in an unspecified manner across physicians and over time does much more damage to the Evans hypothesis than is generally understood. The comment (Evans [7]) that preference functions are not specified by the neoclassical theorist is neither a defense nor accurate. Preference functions are neither completely arbitrary functions nor are they assumed to vary capriciously over time. If that were the case, demand theory would be devoid of all content. Secondly, even if demand theory had been eviscerated by such assumptions, that is no defense for postulating yet another theory equally empty in empirical implications or content. Even if one ignores all the other logical difficulties in the Evans model, the occurrence of an unspecified and even unmeasurable variable N^I in the list of exogenous variables in the supposedly simultaneous equation system (equation (19)) nullifies the entire analysis and every sign prediction, which were all computed under the assumption that N^I is constant. But N^I constant is generally assumed to be false, as even Evans himself admits [7, p. 45], so that this model in reality makes no predictions at all.

The formulation of the price equation introduces yet another aspect of the logical difficulties created by not deriving empirical relationships from basic theoretical premises. In terms of the model used to make sign predictions, price is assumed to be set by the physician in response to variations in $N - N^I$ only. One must then

ask why P is not set so that $N = N^1$; if not, what is the optimizing behavior that leads to a choice of P which leaves $N \neq N^1$? Not only is N^1 assumed to vary capriciously (at least as far as the econometric observer is concerned), but it seems to play no effective role in this model.

A minor difficulty occurs with the variable MD, which is defined as an endogenous variable, but is then used as an exogenous variable in the formulation of equation (19). Further, since all the model's equations were formulated in deterministic form and a considerable number of algebraic manipulations underlay the derivation of (19), it is not at all clear that a stochastic formulation of the model in terms of the "structural equations" would yield an estimable model, nor is it clear that the model is identified; partly because one of the endogenous variables were treated as exogenous in equation (19) and because most of the postulated equations from which (19) is derived are not structural equations in the first place.

Finally, if one ignores all the difficulties discussed above and merely tries to extract a useful concept from the presentation of the model, one conclusion is clear. The model is not a presentation of a simultaneous equation system in static equilibrium, but some form of adjustment, or rather non-adjustment, model. The existence of a situation in which $N \neq N^1$ and $W \neq W^d$ (or $S \neq 0$) implies that the model is not one of equilibrium. However, the formulation of the model does not specify how the model adjusts toward equilibrium either, nor is the nature of the equilibrium specified; one might plausibly assume that equilibrium occurs when $N = N^1$ and $W = W^d$. In this regard, a further logical difficulty might be mentioned: variables defined as functions of the indicators of disequilibria, $N - N^1$ and S, should really be defined in terms of the change in the variables; that is, instead of $P = P((N - N^1), S)$, one should consider $\Delta P = P((N - N^1), S)$, where ΔP indicates the change, or rate of change in prices.

The idea of examining adjustment paths adds a new dimension to the analysis in the previous section. If we conclude from the discussion in this section that the prime objective of Evans model II was to provide a short-run

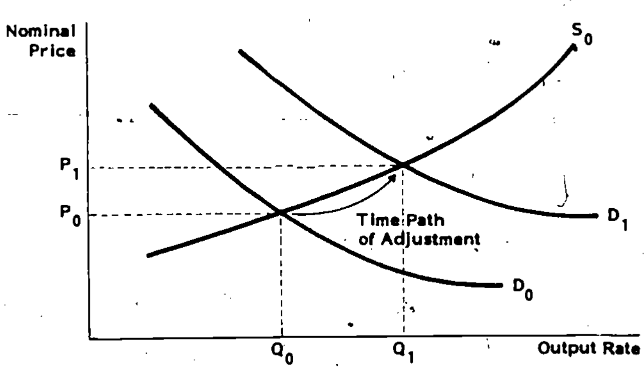


Figure 5.
Illustration of the Process of Short-Run, Adjustment (Price Takers Market)

adjustment model in order to explain the apparent anomalies for the standard model, we should first consider the behavior of feasible short-run adjustment models based on the standard analysis.

To pose the idea is to answer the question. Consider the simplest version of the standard model with individual firm profit maximization in a competitive (price takers) market. A formal model of adjustment is not needed to show that the process of short-run adjustment in the standard model also "explains" the anomalous observations. Consider Figure 5. Suppose for whatever reason market demand is increased from D_0 to D_1, for example due to an increase in population or income. Consider the adjustment path from (Q_0, P_0) to (Q_1, P_1). Immediately after the change, workloads (x_n) and output (y_1) increase, real prices rise (with nominal prices assumed fixed to begin with) due to increased queues, decreased time per visit, etc. In the second stage nominal prices begin to rise, rate of increase in workload (x_n) moderates and then falls as new physicians begin entering the market, thereby lowering the population/physician ratio. During the major portion of the adjustment period, after the initial very short-run reaction to the increase in demand, one observes prices up, population/physician ratios down, individual workloads and output rates down, all of which is compatible with naive interpretations of the observed data.

The importance of this discussion is not to demonstrate that the adjustment process is the explanation of the observed data, but to indicate that insofar as Evans model II purports to explain those observations in terms of a short-run adjustment process, such observations are equally consistent with adjustment paths derived from the standard model.

We may conclude that Evans model II and any other models of a similar nature may be rejected in favor of the standard theory on purely logical criteria; as a class of hypotheses they neither explain observed data nor provide any insight into behavior; they are non-explanations. Perhaps, their resistance to logical argument lies in their emotional appeal and the provision of an apparent justification for further government intervention into the physician/dental market.

Other target income proponents might maintain that the logical arguments addressed above are not relevant to their concept of a "target income." The Evans model was analyzed in detail because, and only because, it appeared to be the most detailed, explicit, and formal development by a leading proponent of the target income concept. The discussion in Part I of this paper indicated that the formal development of one's theories is important, because without such a development it is difficult, if not at times impossible, to check for logical inconsistencies, distinguish structural from semi-reduced form relationships, and be able to evaluate correctly the partial relationships between pairs of

variables embedded in a complex interdependent system.

If some target income concept were to be worthy of empirical testing, its proponents must first be able to demonstrate clearly:

(i) The reconciliation between target income concepts and utility maximization;
(ii) The relationship with modern developments in consumer theory which take into consideration the time and income constraints on use of leisure;
(iii) Why doctors and dentists have "targets" and the rest of us do not (for example, airline pilots probably have even greater market power and and there is precious little evidence pilots have targets).
(iv) Why targets did not apply during the '30's.

PART III. IMPLICATIONS OF THE THEORETICAL ANALYSIS FOR EMPIRICAL RESEARCH

The main conclusion from Part II of this paper is that none of the theories as currently formulated explain the anomalous empirical relationships summarized in Part I. Consequently, the major objectives of this part of the paper are to indicate both how the standard theory might be applicable to the observed data, once what was observed has been interpreted more carefully than in previous studies, and how the prior analysis should be analyzed in order to separate theoretically important findings from statistical artifacts.

In this regard the next section discusses very briefly inferential problems created by specification errors in regression analysis. The following section utilizes the discussion on specification errors in order to comment on some relatively important deficiencies in the current analyses of medical demand and supply data. These deficiencies might prove on further examination to explain the empirical findings seemingly at variance with standard theory. The paper concludes with a set of recommendations for future work in order to resolve to a reasonable level of satisfaction the main issues which were discussed in Parts I and II.

Inferential Problems in the Presence of Specification Errors

In the literature discussed briefly in Part I, there is much discussion about estimation bias in the regression equations because of the presence of a simultaneous equation framework. Unfortunately, in models of the type examined in this paper, this "bias problem" is by far the least important difficulty facing the researcher. Another relatively unimportant issue is the problem of "errors in the variables," wherein the errors are assumed to be non-systematic random variations with zero mean. The crucial issues involve other types of error.

The single most important issue is identification. Further, as will be discussed, identification is a more complex issue than researchers in this area of inquiry have recognized. Indeed, the major difficulty with the presence of the other types of specification error to be mentioned later lies in the implications of such errors for identification.

By now every applied economist is well aware of the necessary conditions for evaluating the identifiability within the simultaneous equation model of any proposed equation; namely that the available amount of information exogenous to the equation under study should be greater than the equation's requirements for such information, i.e., in the simplest of circumstances the number of excluded exogenous variables should be greater than the number of endogenous variables included as regressors. An aspect of this requirement that is often overlooked is that the excluded exogenous variables must not be multicollinear. For example, it is easy to postulate a simple demand/supply model wherein each equation is identified according to the usual necessary conditions, but that the pattern of shifts in the apparently "identifying" exogenous variables are such as to yield, for example, a constant expected value for the market price. Secondly, it is often forgotten that identification is achieved only insofar as the identifying variables are in fact non-trivial variables in the system and that over the observed period the sample variances of the identifying variables are significantly different from zero, where "significance" can be measured in terms of the effect on the conditional means of the endogenous variables.

The second most important specification error problem is that of omitted variables and use of incorrect functional forms for relating the conditional means to the regressors. Since least squares does the "best job" of fitting the regression line to the included data, the criteria of high R^2, or even high $\bar{R}^2$, plausible signs for some coefficients, and high "t" ratios, are inadequate, if not completely useless, indicators of the presence of specification errors (for a more detailed discussion see, for example, [27, 28, 29]). In summary, even seriously misspecified models do not always (nor even frequently) declare themselves in terms of the conventional criteria for a "good regression".

While any researcher can always speculate on a wide variety of potential sources of specification error, the important issue is whether the specification errors are such as to nullify the inferential implications. To put the matter slightly more precisely, are the errors sufficient to produce measurable systematic effects on the inferences to be drawn? Such questions are answered by the use of the specification error tests discussed in [28, 29].

A related problem is the use of proxy variables in a regression analysis, an almost inevitable practice in econometric research. Unfortunately and all too often, researchers using proxy variables do not bother to evaluate the effect of the use of such variables on their regression results, at least in a qualitative sense. A not so subtle aspect of this problem is that if a given proxy

variable is regarded as composed of two component parts, a trend component and a variable specific component, then researchers usually choose the proxy on the basis of the trend component that is common to most of the variables in the system, whereas the regression significance of the variable lies in its specific component. As a consequence, one can often pick entirely inappropriate proxy variables. What one wishes to be able to do, of course, is to pick a proxy variable such that its contribution to the regression *net* of the contribution of the other variables is nearly the same as the *net* contribution under the same circumstances as the unobservable, but appropriate variable. Operationally, what these comments imply, is that if one has some theoretical information about the net behavior of the unobserved variable, then one might find it advantageous to choose proxy variables on the basis of the behavior of the residuals obtained from regressing the proxy against the other regressors. For example, if the unobserved true variable is known to have increased over time relatively more quickly in an early period and relatively less in a later period, and one has a choice between two proxy variables, one which exhibits such behavior and the other which does not, one would be advised to use the former, rather than the latter.

This section can be summarized by saying that before one begins an elaborate search for new hypotheses of economic behavior, it pays to analyze the model very carefully for the presence of specification errors. Evidence of such errors implies not only that the significance of coefficient estimates, or the lack of it, is suspect, but even the signs of supposedly highly signficant (statistically) coefficient estimates as well as the results of tests of hypotheses are also suspect.

Some Particular Difficulties in Estimating Physician Demand/Supply Equations

A few problems in the estimation of physician supply/demand relationships are of particular significance. As mentioned in Part I, two main items of empirical evidence have stimulated the search for non-standard hypotheses: positive partial correlations between price and quantity demanded and a negative partial correlation between population/physician ratios and price. The former difficulty occurs mainly in time series and the latter mainly in cross-sectional work.

One of the more important aspects of medical care that is difficult to measure is technological change. The main problem does not lie in the shift in the production function, but in the change in the medical options faced by consumers as well as the dramatic change in the relative prices of alternative medical services. Thus, over time there appears to have been a sizable shift in the composition of services provided in any representative "basket of services," both in terms of the services provided by each doctor and in terms of the percentage of doctors in various specialties. Even if demand had not shifted in response to income and population over time, any observed (constant weighted) price index would not be a suitable measure of relative changes in physician fees, and plausibly would erroneously indicate a substantial increase in the realtive price of physician fees. For example, if certain ailments previously treated by an attending physician can subsequently be handled by drugs, so that even general practitioners tend to specialize in more resource using services, the observed "average fee per visit" will rise. In addition, if price indices are averaged over general practitioners as well as specialists, a rise in the percentage of specialists demanded (because of a technologically induced decrease in the relative price of specialists) will be evidenced by an *apparent* increase in the index of physician fees.

Another factor missing from the analysis of the three-decade period usually involved in time studies is the induced change in patient-perceived relative prices of various medical services because of differences in the coverage of different medical services by insurance. In general, it would appear that the more expensive treatments are insured to a relatively greater extent, thereby lowering to the patient the cost of high resource use services relative to low resource use services. The change in patient-perceived relative prices of "expensive" to "cheap" medical services would induce patients to increase consumption of the former relative to the latter; for example, under insurance the price of specialists is relatively less to that of general practitioners than without insurance.

A further factor which researchers in the future will have to take into account is the shift in the expected cost of physicians being sued. Insofar as each physician can lower his suit costs by his own actions, risk-averse behavior indicates that the use of diagnostic checking and the solicitation of second opinions will increase, with corresponding increases in the nominal price of physician services.

In terms of cross-sectional studies, a researcher must be careful to allow for the non-homogeneous geographic dispersion of various types of physician. For example, average fees calculated in a predominantly rural area will be less than in a predominantly urban area in that the percentage of general practitioners in the former area is much greater than in the latter. Further, since the more specialized a physician, the greater the geographic extent of the relevant market, and the economically justified fact that it is cheaper for specialists to congregate in centers, a comparison of "urban" and "rural" areas would indicate a negative relationship between the price index and the ratio of population to physicians.

A further contributing factor is that given the restricted geographic size of physician markets, a cross-sectional study will usually involve a comparison across markets separated by transportation costs. Thus, even if physicians can migrate easily across markets, other inputs, such as office space, cannot, and a portion of

ıysician fees includes costs of providing medical rvices from factors other than medical personnel. In ort, while an efficient market hypothesis would ınclude that real physician *wage rates*, w_n in the nguage of Part II, would be constant across regions 1 equilibrium), one would not expect physician *fees* to : equal, but to be lower in low population density areas ıd higher in denser areas.

A final problem, which is perhaps of little practical gnificance, is that in many low population density eas, physicians often receive direct and indirect ıbsidies to practice in the area; thus, the observed ɔminal fee understates the actual opportunity cost of ·oviding medical services in the area.

ecommendations for Further Work

The most effective way to summarize the import of is paper is to indicate the further work needed to solve the current uncertainties about the relevance of ındard theory. There are several issues of interest.

The least important, perhaps, is whether the ıysician and dental markets are competitive or only onopolistically competitive; alternatively stated, are ıysicians and dentists price takers or price searchers? s was seen, this issue separates into two components, e pricing behavior of physician/dental firms and the arket for physicians/dentists. Thus, the questions lative to this topic resolve themselves into these ıestions:

(i) Are individual physicians/dentists price takers, notwithstanding entry limitations by medical/dental societies which are supported by state or federal law?

(ii) Even if the answer to (i) is yes, one can ask whether physicians or dentists are earning rents from medical and dental society restrictions on entry,

(iii) Whether physicians/dental firms are price takers or price searchers, does the efficient physician/dental market hypothesis fail?

ıe discussion in Part II provides the appropriate eoretical framework for beginning the empirical ıalysis of these questions.

The second question, not unrelated to part (iii) of the ·st, is whether the supply-induced demand hypothesis ›lds. This hypothesis was formulated as Evans model but as interpreted by this researcher, in Part II of the ıper. Two procedures were suggested for testing that 'pothesis.

The third set of questions concern the "explanation" the empirical results commented upon in Part I. ıere are two aspects to the approach to this problem. rst, the prior analyses need careful examination for e presence of specification errors and lack of entification. As a part of this approach, the eoretical framework needs to be reconsidered and in me cases elaborated in order to reevaluate more carefully the identification issue. The standard model discussed in Part II supplemented by a development of the market for medical insurance would provide a useful starting point for this analysis.

The second aspect of this approach is to examine new ways in which to pay much closer attention to the effects of technological change and medical insurance on consumer medical service options as well as the significant change in relative prices of medical services, not to mention the increased role of drugs. Cross-sectional studies must be careful not to confound "within market effects" with "between market effects."

By these means some commonly agreed resolution of the disagreements about the nature of physician and dental markets might be achieved.

References

1. Acton JP: Nonmonetary factors in the demand for medical services: Some empirical evidence. *Journal of Political Economy* 83:595-614, 1975
2. Backman J: Professional fees: Factors affecting fee setting in the several professions. *The Journal of Accountancy* 95:554-566, 1953
3. Benham L, Maurizi A, Reder MW: Migration, location and renumeration of medical personnel: Physicians and dentists. *The Review of Economics and Statistics* 50:332-347, 1968
4. Brown DM, Lapan HE: The rising price of physicians' services: A comment. *The Review of Economics and Statistics* 54:101-104, 1972
5. Drhymes PJ, Howrey EP, Hymans SH, Kmenta J, Leamer EE, Quandt RE, Ramsey JB, Shapiro HT, Zarnowitz V: Criteria for evaluation of econometric models. *Annals of Economic and Social Measurement* 1:291-324, 1972
6. Dyckman ZY: *A Study of Physicians' Fees*. Council on Wage and Price Stability, Executive Office of the President, Washington, March 1978
7. Evans RG: *Price Formation in the Market for Physician Services in Canada 1957-1969*. Prices and Income Commission, July 1972
8. Evans RG, Parish EMA, Sully F: Medical productivity, scale effects, and demand generation. *Canadian Journal of Economics* VI:376-393, 1973
9. Evans RG: Supplier-induced demand: Some empirical evidence and implications. In *The Economics of Health and Medical Care* (Perlman M, ed.). London, Macmillan Co., 1974, pp 162-173
10. Evans RG: Beyond the medical marketplace: Expenditure, utilization, and pricing of insured health care in Canada. In: *The Role of Health Insurance in the Health Services Sector* (Rosett RN, ed.). New York, National Bureau of Economic Research, 1976, pp 437-492
11. Feldstein MS: The rising price of physicians' services. *The Review of Economics and Statistics*. LII: 121-133, 1970

12. Feldstein MS: The rising price of physicians' services: A reply. *The Review of Economics and Statistics* 54:105-107, 1972
13. Feldstein M, Friedman B: The effect of national health insurance on the price and quantity of medical care. In: *The Role of Health Insurance in the Health Services Sector* (Rosett RN, ed.). New York, National Bureau of Economic Research, 1976, pp 505-529
14. Feldstein PJ: *Financing Dental Care: An Economic Analysis*. Massachusetts, Lexington Books, 1973
15. Frech HE, Ginsburg PE: Physician pricing: monopolistic or competitive: comment. *Southern Economic Journal* 38:573-577, 1972
16. Fuchs VR, Kramer MJ: *Determinants of Expenditures for Physicians' Services in the United States, 1948-68*. Mimeo., National Bureau of Economic Research, December 1972.
17. Guzick DS, Jahiel RI: Distribution of private practice offices of physicians with specified characteristics among urban neighborhoods. *Medical Care* 14:469-488, 1976
18. Holen A: *The Economics of Dental Licensing*. The Public Research Institute of the Center for Naval Analyses, November 1977
19. Kessel RA: Price discrimination in medicine. *Journal of Law and Economics* 1:20-53, 1958
20. Kushman JE, Scheffler RM: Pricing health services: Verification of a monopoly pricing model for dentists. *Journal of Human Resources* 13:402-415, 1978
21. Newhouse JP: A model of physician pricing. *Southern Economic Journal* 37:174-183, 1970
22. Newhouse JP, Sloan FA: Physician pricing: Monopolistic or competitive: Reply. *Southern Economic Journal* 38:577-580, 1972
23. Newhouse JP, Phelps CE: New estimates of price and income elasticities of medical care services. In: *The Role of Health Insurance in the Health Services Sector* (Rosett RN, ed.). New York, National Bureau of Economic Research, 1976, pp 261-313
24. Nordquist G, Wu SY: The joint demand for health insurance and preventive medicine. In: *The Role of Health Insurance in the Health Services Sector* (Rosett RN, ed.). New York, National Bureau of Economic Research, 1976, pp 35-65
25. Olsen EO: Utility and profit maximization by an owner-manager. *Southern Economic Journal* 39: 389-395, 1973
26. Pashigian BP: Occupational licensing and the interstate mobility of professionals. Mimeo, National Bureau of Economic Research, February 1978
27. Ramsey JB: Tests for specification errors in classical linear least-squares regression analysis. *The Journal of the Royal Statistical Society* 31:350-371, 1969
28. Ramsey JB: Models, specification error, and inference: A discussion of some problems in econometric methodology. *Bulletin of the Oxford University Institute of Economics and Statistics* 32:301-318, 1970
29. Ramsey JB: Classical model selection through specification error tests. In: *Frontiers in Econometrics*. New York and London, Academic Press, Inc., 1973, pp 13-47
30. Reinhardt U: A production function for physician services. *The Review of Economics and Statistics* 54:55-66, 1972
31. Scheffler RM: A production function for dental services: Estimation and economic implications, *Southern Economic Journal* 44:25-35, 1977
32. Sloan FA, Steinwald B: The role of health insurance in the physicians' services market. *Inquiry* 12:275-299, 1975
33. Sloan FA: Physician fee inflation: Evidence from the late 1960s. In: *The Role of Health Insurance in the Health Services Sector* (Rosett RN, ed.). New York, National Bureau of Economic Research, 1976, pp 321-354
34. Sloan FA: Real returns to medical education: A comment. *The Journal of Human Resources* XI: 118-126, 1976
35. Sloan FA, Lorant JH: The allocation of physicians' services: Evidence on length-of-visit. *Quarterly Review of Economics and Business* 16:85-103, 1976
36. Sloan FA: Access to medical care and the local supply of physicians. *Medical Care* XV:338-346, 1977
37. Sloan FA, Lorant JH: The role of waiting time: Evidence from physicians' practices. *Journal of Business* 50:486-507, 1977
38. Sloan F, Feldman R: Monopolistic elements in the market for physicians' services, a paper presented at a conference on "Competition in the Health Care Sector: Past, Present, and Future," June 1977
39. Sloan FA, Steinwald B: Physician participation in health insurance plans: Evidence on Blue Shield. *Journal of Human Resources* 13:237-263, Spring 1978
40. Steinwald B, Sloan FA: Determinants of physicians' fees. *Journal of Business* 47:493-511, 1974

A Formal Model of Target Income Pricing with Supplier-Induced Demand

Charles S. Roehrig, Ph.D.
Policy Analysis, Inc.

ABSTRACT □ The purpose of this paper is to provide a mathematical model of the market for physicians' services which includes precisely defined concepts of supplier-induced demand (demand creation) and target income pricing, and which can contribute to the proper interpretation of empirical evidence pertaining to physician behavior.

The physician is assumed to be a utility maximizer, with utility taken to be a function of income, leisure, and the amount of created demand; it is assumed that there is a disutility to the physician in creating demand. Target income pricing can be viewed either as an external constraint on prices or as the result of peculiarities in the utility function. In the absence of external constraints on pricing, the model assumes monopolistic competition with a parameter that can be set to reflect pricing anywhere between perfect competition and monopoly.

Equations are derived for the effect of changes in basic parameters (the population to physician ratio, type and extent of health insurance coverage, costs of malpractice insurance, etc.) on the price of care and the amount of created demand.

BACKGROUND

As economists have begun to investigate the market for physicians' services, evidence has accumulated which suggests that traditional economic models do not apply. Attention has centered upon the extent to which physicians can influence the patient's perceived need for care, and upon empirical evidence that increasing the number of physicians in an area might actually cause increases in the prices charged for care.

Specifically, it is thought that as the market for phsyicians' services deteriorates through increased supply, physician incomes are maintained by increasing prices by the use of their power to "artificially" expand demand. This ability to maintain income levels in the face of declining demand implies that incomes were not maximized in the first place. Instead, there is some "target" income, below the maximum level, which physicians seek to maintain.

The purpose of this paper is to provide a mathematical model of the market for physicians' services which includes precisely defined concepts of supplier-induced demand (demand creation) and target income pricing, and which can contribute to the proper interpretation of empirical evidence pertaining to physician behavior. There will be no discussion of the characteristics of physician practices which relate to the actual process of demand creation and target income pricing. It is the implication of this hypothesized behavior in terms of market responses (prices, utilization, physician incomes, etc.) to variations in underlying conditions (physician-population ratios, health care insurance coverage, etc.) which is the central issue here.

GENERAL DESCRIPTION OF THE MODEL

The provider in the model is a single physician with output directly proportional to the hours worked. The demand for care is divided into two components: patient-initiated demand and return visits initiated by the physician. Patient-initiated demand is a downward sloping function of the price paid by the patient, with

the possibility that because of insurance coverage this price is less than the price charged by the physician. Physician-initiated demand is itself divided into two components, referred to as the "normal" component and the "created" component. The normal component is that level of return visits that would occur if the physician provided the patient with complete and accurate information as to the benefits to be gained from additional visits. Any deviation from this normal level is referred to as created demand. The physician is assumed to be a utility maximizer, with utility taken to be a function of income, leisure, and the amount of created demand; it is assumed that there is a disutility to the physician in creating demand.[9]

The concept of target income pricing is not clearly defined in the literature and is therefore not conducive to a single mathematical interpretation. At one extreme it could be defined as pricing such that physicians can earn the "target" income, regardless of the workload in their geographic area. This would imply higher prices in areas with low workloads. At the other extreme, prices could be set to the same level over all areas with the level determined such that physicians with a full workload earn the target income. In the model, it will be assumed that prices are a function of workload with the parameters of that function limited by these two extremes.

BASIC EQUATIONS OF THE MODEL

Demand Equations

Patient-initiated Demand

Patient-initiated demand for care arises from the patient's perceived need for care and should be negatively related to the price he pays. In equation form:

$$Q'_p = A_p + B_p P_p \qquad (1)$$

where Q'_p is the average per capita patient initiated demand when the price paid by the patient is P_p and B_p is negative.

Physician-initiated Demand

Normal expansion. Once the patient has initiated a visit to the physician, the possibility arises that return visits will be agreed upon. It is here that the physician can exercise power over the demand for his services. If he gives the patient complete and accurate information regarding the costs and benefits likely to result from return visits, there should be some expansion of patient-initiated demand which will be referred to as "normal." The normal per capita demand for care is given in equation form by multiplying patient-initiated demand as represented in (1) by the normal expansion factor K_N:

$$Q'_N = A' + B'P_p \qquad (2)$$

where $A' = K_N A_p$ and $B' = K_N B_p$.

Demand Creation

Actual demand for care will vary from the normal level to the extent that the physician does not provide complete and accurate information to the patient. This will occur when the physician's presentation of options to the patient is influenced by the desire to affect his workload in a particular way. Most commonly this would involve encouraging the patient to receive care because of a desire to expand workload rather than because it would be beneficial to the patient. In cases where workload is very high, it is possible that constraints on price increases could lead to physicians cutting their demand by discouraging care which would otherwise be advised. Thus, the model allows for both positive and negative demand creation. Letting Q'_c represent the average per capita created demand, the average per capita demand including all components is given by

$$Q' = A' + B'P_p + Q'_c \qquad (3)$$

Insurance

It is assumed that the patient pays the first D dollars of the visit price P plus some faction c of the remainder. Thus the price to the patient is given by:

$$P_p = D + cP \qquad D < P \qquad (4)$$

$$P_p = P \qquad D > P \qquad (4a)$$

To simplify notation, it will be assumed that the deductible D is always less than P so that (4) is the appropriate equation. The demand equation can then be written as a function of the price received by the physician (P):

$$Q' = A + BP + Q'_c \qquad (5)$$

where $A = A' + B'(1-c)D$ and $B = cB'$.

Demand Per Physician

The per capita demand equation can be converted to an equation giving the average demand per physician (given that he charges the market price) by multiplying by the ratio of population to the number of physicians:

$$Q = nA + nBP + nQ'_c = Q_N + Q_c \qquad (6)$$

where $Q_N = nQ'_N$ and $Q_c = nQ'_c$.

Production, Cost, and Income Equations

Output is taken to be directly proportional to hours worked by the physician, with marginal cost constant. While this is an over-simplified view of production, it appears that more realistic equations would lead to more complex mathematics without significantly contributing to the model's ability to analyze questions concerning target income and demand creation. The basic production equation is

$$Q = bPH \qquad (7)$$

where PH is physician hours. Because the physician's leisure time enters into the utility function, it is useful to

convert this equation to a function of leisure. The relationship between work and leisure is given by

$$PH = d - Le \quad (8)$$

where Le is leisure and d is the total time available for work or leisure. Combining (7) and (8) gives output as a function of leisure:

$$Q = a - bLe \quad (9)$$

where $a = bd$.

The total cost of producing Q is given by

$$TC = FC + MC_0Q \quad (10)$$

where FC is fixed cost and MC_0 is marginal cost which does not vary with Q. The physician's total net income is.

$$TY = PQ - FC - MC_0Q \quad (11)$$

Utility Functions

The physician is assumed to be a utility maximizer, with utility a function of income, leisure, and the extent to which the physician engages in demand creation. Because the disutility of demand creation should not depend upon income or leisure, there are two separate utility functions in the model; one for the utility of combinations of income and leisure, and one for the disutility of demand creation.

The Income-Leisure Utility Function

This utility function is represented in the model as

$$U = U(TY,Le) \quad (12)$$

It is assumed that the marginal utility of income declines as income increases (leisure fixed), and that the marginal utility of leisure declines as leisure increases (income fixed), with the marginal utilities always positive. Using partial derivatives, this implies

$$U_1 \text{ and } U_2 > 0$$
$$U_{11} \text{ and } U_{22} < 0$$

Finally, it is sensible to assume that the marginal utility of income inçreseases with increases in leisure (income fixed) and that the marginal utility of leisure increases with increases in income (leisure fixed). In terms of partial derivatives

$$U_{12} \text{ and } U_{21} > 0$$

The Disutility of Demand Creation

The disutility of demand creation is taken to be a function of the ratio of created demand to normal demand, where normal demand is the level of demand resulting when there is no demand creation. This is most easily understood using visits as the measure of care. Each normal visit provides the physician with an opportunity to create extra visits and therefore the more normal visits he services, the easier it is to create visits. For an individual physician, the number of created visits per normal visit is Q_c/Q_N. On the assumption that each additional created visit per normal visit is more difficult to justify than the one before, the disutility per normal visit is $-\ a(Q_c/Q_N)^2$. Then the total disutility per physician is this multiplied by the number of normal visits per physician:

$$DU = Q_N(-a)(Q_c/Q_N)^2 = -aQ_c^2/Q_N \quad (13)$$

Note that this treats positive and negative demand creation as equivalent in terms of disutility. A more realistic assumption would have a depend upon the sign of Q_c, but for analyses in which the sign of Q_c does not change (i.e., Q_c always positive or always negative) this issue becomes irrelevant.

Monopolistic Competition

The price P appearing in the previous equations is the "average" or "market" price of care and is appropriate because the model is based on the behavior of the "average" physician. In equilibrium, this average physician charges the market price because doing otherwise would not increase his total utility. If this were not so, the average physician would change his price and thus the market price would change. To determine what the market price will be in equilibrium, it is necessary to determine the conditions under which deviations from the market price are undesirable to the physician.

Given the market price P, suppose the physician charges price P* not necessarily equal to P. The effect upon the demand for his services depends upon a number of factors yet to be specified. If all other physicians follow suit (resulting in the market price changing to P*), or if the physician has monopolistic control over his share of the market, equation (6) will accurately describe the demand for his services. To the extent that other physicians do not follow his price change and to the extent that he does not have monopolistic control (note that if he does have monopolistic control, other physicians would have no reason to follow his price changes), equation (6) becomes inaccurate. The following equation can be used to present the various responses in physician demand when he charged other than the market price:

$$\begin{aligned} Q^* &= Q_N + Q_c + ngB(P^* - P) \\ &= nA + nBP(1 - g) + ngBP^* \\ &\quad + Q_c, \ g \geq 1 \end{aligned} \quad (14)$$

If $g = 1$, this reduces to equation (6) which is the case of pure monopoly. As g becomes infinite, this represents a perfectly competitive situation in which the physician has as much demand as he wishes to service when he charges at or below the market price, and no demand at all if he charges anything above the market price. Thus values of g between 1.0 and plus infinity represent market situations, between monopoly and competition.

This is depicted graphically in Figure 1 where

$$Q^*_N = Q_N + ngB(P^* - P)$$

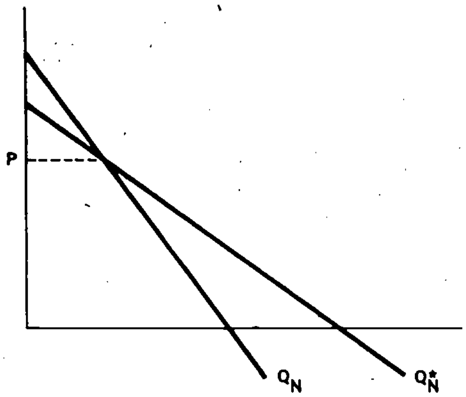


Figure 1

UTILITY MAXIMIZATION

Maximization Without Target Income Constraints

Given the market price P, the current model can be solved for the price P* which maximizes the average physician's total utility. If this utility-maximizing price is not equal to the market price (i.e., if $P^* \neq P$), then the market cannot be in equilibrium because on the average, physicians will adjust their prices towards P*. Therefore, the equilibrium market price P and the amount of created demand (Q_c) per physician can be found by first maximizing the physician's utility given P, and then setting P* equal to P in the solution equations.

Under the assumption of monopolistic competition, the equations constituting the model are:

$$Q^* = nA + nBP(1 - g) + ngBP^* + Q_c \quad (14)$$

$$Le^* = (a - Q^*)/b \quad (15)$$

$$TY^* = P^*Q^* - MC_0Q^* - FC \quad (16)$$

where the function to be maximized is

$$\Phi = U(TY^*, Le^*) + DU^*$$
$$= U(TY^*, Le^*) - aQ_c^2/Q^*_N \quad (17)$$

The physician attempts to maximize his total utility as represented in (17) through adjustments in his price P* and in the amount of created demand Q_c, given the market price P. Therefore, the maximizing solution is found by setting the partial derivatives of Φ with respect to P* and Q_c equal to zero.

$$\partial\Phi/\partial P^* = U_1\partial TY^*/dP^* + U_2\partial Le^*/\partial P^* + \partial DU^*/\partial P^* = 0 \quad (18)$$

$$\partial\Phi/\partial Q_c = U_1\partial TY^*/dQ_c + U_2\partial Le^*/\partial Q_c + \partial DU^*/\partial Q_c = 0 \quad (19)$$

where

$$\partial TY^*/\partial P^* = (P^* - MC_0)ngB + Q^* \quad (20)$$

$$\partial Le^*/\partial P^* = -ngB/b \quad (21)$$

$$\partial DU^*/\partial P^* = a\,ngBQ_c^2/Q_N^{*2} \quad (22)$$

$$\partial TY^*/\partial Q_c = P^* - MC_0 \quad (23)$$

$$\partial Le^*/\partial Q_c = -1/b \quad (24)$$

$$\partial DU^*/\partial Q = -2aQ_c/Q^*_N \quad (25)$$

The second order conditions to assure that this is a maximum are:

$$\frac{\partial^2\Phi}{\partial P^{*2}} < 0 \quad (26)$$

and

$$\left(\frac{\partial^2\Phi}{\partial P^{*2}}\right)\left(\frac{\partial^2\Phi}{\partial Q_c^2}\right) - \left(\frac{\partial^2\Phi}{\partial P^*\partial Q_c}\right)^2 > 0 \quad (27)$$

Economic Implications

Suppose that the system is in equilibrium so that equations (18), (19), (26), and (27) are satisfied and P* = P, and suppose that there is then a shift in the underlying conditions. The purpose of this section is to derive equations for the effect of changes in underlying conditions upon the equilibrium values of P, Q_c, and the remaining quantities determined by P and Q_c (Q, TY, Le, etc.). For example, suppose that the underlying condition that changes is the population-physician ratio. Mathematically, this means solving for dP/dn and dQ_c/dn. As another example, consider a change in the cost of malpractice insurance. Since this is basically a fixed cost, the effect of changes in malpractice premiums is given by dP/dFC and dQ_c/dFC.

For the purposes of illustration, consider a small change in n (the population-physician ratio). Taking the total differentials of (18) and (19) yields:

$$\frac{\partial^2\Phi}{\partial P^{*2}}dP^* + \frac{\partial^2\Phi}{\partial P^*\partial Q_c}dQ_c + \frac{\partial^2\Phi}{\partial P^*dn}dn = 0 \quad (28)$$

$$\frac{\partial^2\Phi}{\partial Q_c^2}dQ_c + \frac{\partial^2\Phi}{\partial Q_c\partial P^*}dP^* + \frac{\partial^2\Phi}{\partial Q_c\partial n}dn = 0 \quad (29)$$

and Cramer's rule can then be used to solve for $\frac{dP^*}{dn}$ and $\frac{dQ_c}{dn}$. Setting P* = P in the resulting solutions yields:

$$\frac{dP}{dn} = \frac{\frac{\partial^2\Phi}{\partial Q_c\partial P^*}\cdot\frac{\partial^2\Phi}{\partial P^*\partial n} - \frac{\partial^2\Phi}{\partial P^{*2}}\cdot\frac{\partial^2\Phi}{\partial Q_c\partial n}}{\frac{\partial^2\Phi}{\partial P^{*2}}\frac{\partial^2\Phi}{\partial Q_c^2} - \left(\frac{\partial^2\Phi}{\partial P^*\partial Q_c}\right)^2} \quad (30)$$

and

$$\frac{DQ_c}{dn} = \frac{\frac{\partial^2\Phi}{\partial Q_c\partial n}\cdot\frac{\partial^2\Phi}{\partial P^*\partial Q_c} - \frac{\partial^2\Phi}{\partial P^*\partial n}\cdot\frac{\partial^2\Phi}{\partial Q_c^2}}{\frac{\partial^2\Phi}{\partial P^{*2}}\frac{\partial^2\Phi}{\partial Q^2} - \left(\frac{\partial^2\Phi}{\partial P^*\partial Q_c}\right)^2} \quad (31)$$

Note that the denominator in (30) must be positive because of the second order conditions (given in (27)) necessary to insure that the solution is a maximum. This means that the signs of dP/dn and dQ_c/dn are determined by the signs of the numerators.

Equations for the effects on P and Q_c of changes in other underlying conditions can be derived in a similar fashion. In each case, the sign of the derivative is determined by the sign of the numerator. Thus the determination of conditions under which the numerator has a particular sign will provide insight into empirical estimates of these derivatives.

It has been found that the price of physicians' services is often positively correlated with the number of physicians per capita in an area, using changes in the population-physician ratio as an example. This finding implies that $\frac{dP}{dn}$ is negative (n is the inverse of physicians per capita) and is cited as evidence of target income pricing. If conditions exist under which the numerator of (30) is negative, these conditions will constitute an alternative explanation of the empirical result because target income pricing is not built into this version of the model.

Maximization with Target Income Constraints

Let us now suppose that the utility maximization process is constrained by target income pricing, which is interpreted as setting prices such that incomes obey some desired relationship in workload. In equation form:

$$P = P_0 + f(Le - Le_0) \tag{32}$$

where f is some function with a non-negative first derivative f'.

Viewed cross-sectionally, this implies that the lower the workload in an area, the higher the prices. f' is assumed to be non-negative but small enough that in no case does income increase with leisure. Thus the two extremes are 1) prices equal over all areas (f' ≡ 0) and 2) incomes equal over all areas.

This price constraint becomes relevant when it is exceeded by the unconstrained utility maximizing price. When this is the case, in the neighborhood of the constrained solution, physicians view their fees as fixed and the utility-maximizing conditions are given by (19) which sets Q_c to maximize utility for any given price P*. In equilibrium, P* = P and P is determined by the target income constraint in (32). Thus the equations determining this point of equilibrium are (19) with P* = P and (32). Solving (19) and setting P* = P yields:

$$Q_c = [U_1(P - MC_0) - U_2/b]\frac{Q_N}{2\alpha} \tag{33}$$

Note that

$$\frac{\partial^2 \Phi}{\partial Q_c^2} = \frac{-2\alpha}{\partial Q_N^*} < 0$$

which means that for all positive values of Q_N^* the solution is a maximum.

Economic Implications

As with the unconstrained solution, these equations can be solved for the effects of changes in the underlying conditions upon prices, utilization, etc., by taking total differentials and using Cramer's rule. For notational convenience, let us rewrite (32) and (33) as

$$G = P - P_0 - f(Le - Le_0)$$

$$H = [U_1(P - MC_0) - U_2/b] - \frac{2\alpha Q_c}{Q_N}$$

Then, for example:

$$\frac{dP}{dn} = \frac{-\frac{\partial G}{\partial P} \cdot \frac{\partial H}{\partial n} + \frac{\partial H}{\partial P} \cdot \frac{\partial G}{\partial n}}{\frac{\partial G}{\partial P} \cdot \frac{\partial H}{\partial Q_c} - \frac{\partial H}{\partial P} \cdot \frac{\partial G}{\partial Q_c}} \tag{34}$$

$$\frac{dQ_c}{dn} = \frac{-\frac{\partial G}{\partial n} \cdot \frac{\partial H}{\partial Q_c} + \frac{\partial H}{\partial n} \cdot \frac{\partial G}{\partial Q_c}}{\frac{\partial G}{\partial P} \cdot \frac{\partial H}{\partial Q_c} - \frac{\partial H}{\partial P} \cdot \frac{\partial G}{\partial Q_c}} \tag{35}$$

Note that the second order condition no longer determines the sign of the denominator as it did in the unconstrained solution.

CONCLUSIONS

The ultimate purpose of this research effort was to determine the relationships between the parameters of the model and the signs of certain derivatives that relate to the target income hypothesis and the concept of demand creation. This information could then be used to interpret empirical results as to the signs of these derivatives by identifying the set of circumstances (i.e., parameter characteristics) under which the observed sign should be expected.

The process of identifying the determinants of the signs of these derivatives requires a large amount of creativity both in terms of algebraic manipulation and in terms of making simplifying assumptions under which definite results are both possible *and* interesting. While the research time invested in this area produced progress in simplifying the overall problem, the results were not sufficiently complete to be reportable. Therefore, the final conclusion of this paper awaits further research.

The Effect of Provider Supply on Price*

Mark V. Pauly, Ph.D.
Mark A. Satterthwaite, Ph.D.
Northwestern University

ABSTRACT ☐ This paper develops the theory and provides some empirical tests of an alternative to the target income model to explain why price levels for medical and other professional services can be positively related to the supply or providers of those services. This alternative theory, called the "increasing monopoly" theory, posits that firm-level demand curves may become less elastic as the number of providers in a market area increases. As a consequence, the income-maximizing price may well increase.

The primary reason why firm-level demand curves may become less elastic is that consumers will tend to have less accurate information about the price and quality of any provider the more providers there are. A model in which consumers obtain information from their friends on quality levels of various providers is sketched, and it is shown that, as the number of providers increases, the average number of friends who use each provider, and consequently the average level of information, will tend to decrease.

A framework for comparative tests of the target income and increasing monopoly theories is constructed. In both theories, the measure of physician stock is assumed to be endogenous, so that 2SLS estimation methods will be used. In contrast to the target income theory, the increasing monopoly makes price a function of the *number* of physicians in the market area rather than the level of the physician-population ratio. In addition, the increasing monopoly theory indicates that variables that are related to the extent of information flow in a community, such as the number of people who moved, should be related to price.

Empirical tests are provided using data on primary physician fees in 100 large metropolitan areas from the Mathematica telephone survey. Results are consistent with the increasing monopoly theory, in that the number of physicians in the market area and measures of information flow are significant. A target income specification provides a less adequate explanation of prices.

INTRODUCTION

When supply increases, price will fall. This prediction, an object of faith and of fact in most economic markets, may not hold in the case of physician and other professional services. The zero order correlation between physicians per capita and various measures of physician fees tends invariably to be positive. Multivariate regression analyses usually also find a positive and significant relationship (Fuchs and Kramer [3], Newhouse [7], Huang and Koropecky [6], although there is a notable exception (Sloan [11]).

The explanation for these seemingly perverse results is usually a variant of the "target income" theory. In its simplest form, this theory postulates that professionals in an area have a target income to which they aspire. When supply increases, each provider prevents his income from falling by "creating" demand for his

*The work upon which this report is based has been supported in part by grants from the Center for Health Services and Policy Research at Northwestern University and the Robert Wood Johnson Foundation. Tryfon Beazoglou lent invaluable service in assembling the data used in this paper.

services and by increasing his price along that new demand curve. An extended version of the theory, in Evans [2], drops the notion that there is a single target income, but hypothesizes that physicians are willing to substitute money income for diagnostic accuracy to some extent when supply increases. When income falls physicians produce less accuracy, shift demand curves out, and possibly increase price. As indicated by Sloan and Feldman [12] and as emphasized by Reinhardt [9], this extended target income theory is compatible with literally any relationship between physician supply and price. On the other hand, orthodox economic theory, in the sense of profit or income maximization, is usually held to imply that price must fall as supply expands.

In this paper we wish to discuss a model — what we call the "increasing monopoly" model — in which a positive relationship between the price of primary care physician's services and supply is consistent with orthodox income or income-leisure utility maximization under the constraint of a consumer demand curve. No recourse to a "demand creation" explanation is required. The model has been described in detail elsewhere by Satterthwaite [10]. The primary goal of this paper is to discuss an empirical application of the model. We begin with the presentation of a simplified (and slightly different) version of the intuition underlying the increasing monopoly model. We then develop some suggestions for empirical measurement of the concepts of the model. Next, we expand the empirical specification to consider the case of long-run equilibrium, in which provider (physician or dentist) stock is endogenous, along with several other important variables. Finally, we adapt both the increasing monopoly model and the target income model for use with some empirical data on primary care physician prices, and present some empirical estimates for each model.

The results we obtain are striking. For the data analyzed (a cross section of primary care physician fees in 92 large Standard Metropolitan Statistical Areas [SMSA] in the early 1970s), the results are consistent with the predictions of the increasing monopoly model and inconsistent with the target income model. These results suggest that, contrary to the current conventional widsom of consumer powerlessness within the medical care market, consumers, through their choices of which physician to patronize, significantly constrain the pricing decisions of primary care physicians. In particular, within those SMSAs where information about competing primary care physicians appears to be relatively good, the prices the physicians charge tend to be relatively low.

THE EFFECT OF THE NUMBER OF SELLERS ON PRICE IN A MARKET WITH UNCERTAIN QUALITY

Medical care is a service about which consumers have substantial uncertainty. They appear to be uncertain about both price and quality. In a previous paper Satterthwaite [10] showed that, when there is uncertainty about the quality of care provided by various sellers, it is possible for an increase in the number of sellers to result in a price increase. Two ideas underlie this conclusion. First, the market for seller's services is appropriately modeled as monopolistic competition. Second, and of fundamental importance, consumer information, and consequently the degree of certainty about other sellers' quality levels, may decrease as the number of sellers increases. As a result, an increase in the number of sellers may cause the demand curves facing individual sellers to become less elastic and cause each seller's equilibrium price to rise. In what follows we develop a simplified, less formal version of this model and derive similar conclusions.

Consider an individual with a von Neumann-Morgenstern utility function

$$U = U(X,Z).$$

Assume that the good X can be purchased at a price of unity, and that good Z can only be purchased in a "bundle" at a price P. Each buyer buys one "unit" or bundle of the good, but the quality of the good is variable. We capture the notion of variable quality by supposing that different sellers provide varying amounts of Z in their bundles. Thus, if seller i charges P^i and provides a bundle with Z^i units in it, and if consumer a is currently buying from seller i; the consumer's utility is given by

$$U^{ai} = U^a(Y\text{-}P^i, Z^i).$$

Assume that each consumer knows the level of Z provided by his current seller with certainty. Further assume that for a seller j other than his current seller, the consumer knows P^j with certainty, but knows Z^j imperfectly, i.e., has a subjective probability distribution $F^a(Z^j)$ over Z^j. An equilibrium occurs when, for every consumer a, every current seller i, and for every other seller j:

$$U^{ai} \geq \int U^a(Y\text{-}P^j, Z^j)dF(Z^j),$$

i.e., the utility of continuing to patronize seller i exceeds the expected utility of patronizing any other seller j.

The subjective probability distributions $F^a(Z^j)$ depend on the information that person a has about any seller j. It is assumed that, for any seller j, EU^{ai} is greater, the more information that a has about[1] j. That is, information reduces risk-averse person a's uncertainty and so, on average, increases his expected utility if he were to purchase from j. Of course, it may also change his ranking of various sellers, but that is not critical here.

What determines the extent of consumer information about any seller? Our assumption is that the consumer gets information about physician-sellers by asking his friends about their experience with the several sellers within their community. For the simplest possible

[1]Strictly speaking, this assumption requires that information leave the mean of Z unchanged while reducing its "spread." Since information can take many forms, this assumption may not hold exactly.

model, assume that each consumer contacts N friends to obtain information about the quality levels of various providers, and each friend only reports on his experience with his current physician. If we define a unit of information to be the information a single friend reports about his physician, then, on average, the consumer obtains from his N friends N/M units of information about each of the M providers.

The question of interest is: what are the comparative static effects of increasing M on the price of providers' services? Clearly an increase in M decreases N/M and consequently reduces the average amount of information the consumer has about each provider. Consider an individual who, prior to the increase in M, was just on the margin between staying with his current provider i and switching to his next best alternative, seller j. Prior to the increase in M, any finite rise in his current provider's price would have caused this consumer to switch. After the increase in provider stock, however, this individual will have a strict preference for his current seller compared to any alternative. Specifically, let $F_0^a(Z_j)$ represent the consumer's initial subjective probability distribution, and $F_1^a(Z_j)$ the distribution after the increase in M. On average, for any seller j other than the consumer's current seller, the consumer's information levels will have changed in such a way that

$$\int U^a(Y-P^j, Z^j)dF_0^a(Z_j) > \int U^a(Y-P^j, Z^j)dF_1^a(Z_j).$$

As a result, after the increase in M, the seller i's price can rise and yet the consumer will not switch to his next most preferred alternative, seller j; the demand facing the seller will therefore be less price elastic. Income-maximizing providers will be able to respond to this less elastic demand by raising their fees. The effect of increasing stock can be to produce a new monopolistically competitive equilibrium with higher prices.

This conclusion is robust under changes in the assumptions. For example, if the consumer is permitted to vary the number of friends he asks for advice, the same result will follow. If the "cost" of asking a friend about his experiences given, then the effect of increasing M is to reduce the average amount of information that a given number of friends' reports can produce about each seller's quality. This, in effect, raises the price of obtaining a unit of information about any particular seller. This increased cost of obtaining information will, after M increases, lead the consumer to ask fewer friends, and thus lead to a lower equilibrium level of consumer information. The argument then proceeds exactly as above to the conclusion that an increase in M can cause an increase in equilibrium price.

TOWARDS EMPIRICAL ESTIMATION

The foregoing model, which we shall call the "increasing monopoly" model, makes average provider price a decreasing function of N/M or, more conveniently, an increasing function of M/N. Note that M/N is just the number of sellers per information source that a buyer uses. In empirical application to the markets for physicians' and dentists' services, a reasonable empirical approximation of this theoretical concept would be the number of physicians or dentists in the buyer's market area divided by the number of friends that he has available to consult.

Before we discuss what empirically measurable variables might correspond to these concepts, it is useful to contrast this empirical implication of the increasing monopoly model with that of the target income or extended utility maximization model. The latter model is consistent with a positive relationship between providers per capita and price. (It is also consistent with any other relationship, as Reinhardt [9] has noted.) With demand held constant, increases in the number of providers will reduce provider workloads, which will in turn reduce money income unless price is raised. Note that this qualitative prediction is independent of the number of providers in the market area; it holds equally well if the number of physicians or dentists in an area increases from one to two as if it increases from 1000 to 2000, population held constant. Only the ratio of sellers to population is relevant.

The increasing monopoly approach, on the other hand, makes price a function of the *number* of sellers, and not just of their relationship to population. In this view, if there are 4 sellers in an isolated town, price will be low compared to an area in which there are many physicians, even if both areas have the same provider-population ratio. With market area population held constant, of course, the two measures — number of sellers and sellers per capita — will be perfectly correlated. Even if market area population varies over observations, there is still a tendency for correlation.

Empirical testing therefore may yield results supporting both models because the critical variables — number of sellers in the market area and sellers per capita — are correlated. They are not perfectly correlated, however, so if both models "work" in the sense that the number of sellers and the number of sellers per capita have statistically significant coefficients of the predicted sign, then it may still be possible to determine which works *better*. "Better" here is defined in terms of explanatory power and in terms of the significance of other variables that one model suggests and the other does not.

There is an important additional difference in the way in which each of the two models could be tested empirically. The sellers per capita variable required by the target income model is directly measurable, but the sellers per information source variable (M/N) of the "increasing monopoly" model is not directly observable with available data. The remainder of this section, therefore, proposes a set of proxy variables that can be used in the estimation procedure as substitutes for

M/N. Consider first proxies for the numerator, M, which is the number of sellers in the average consumer's market area. In all but smaller, isolated cities, a consumer's market area will be only a fraction of the geographical area of the community in which he is located. For example, a west-sider only considers west side and down-town physicians, but not east side physicians. Consequently, except for small cities, the total number of sellers in the area will not be a suitable measure of M. The simplest case for which a proxy measure for M could be obtained would be if travel cost per mile is the same in all communities in the area, if consumers seek only sellers who can be reached at a given travel cost from their homes, and if in each community population is distributed at uniform density over the community's land area. These conditions would imply that the number of sellers in a consumer's market area is proportional to the number of sellers per unit area in the community.

This simple case is not likely to correspond to reality. Travel time per mile does vary from city to city, and sellers in the downtown area (if the city has a defined downtown) may be accessible to all consumers while suburban sellers are not. For example, consumers in a congested city with high travel time and no well defined downtown would tend to choose their physician from a small market area. Variables that appear, *a priori,* to be related to travel time and community geography (existence of a downtown) are the fraction of the workforce that takes public transportation to work and the population density. These variables, when combined with measures of the number of sellers per unit area, are reasonable candidates to be proxies for the M term of the unobservable variable M/N.

Consider now the denominator, N, which is the number of information sources a consumer is likely to consult when searching for a seller. Given a market area, the number of information sources can be approximated by the number of friends or contacts an average consumer has and the frequency of his contact with those friends. Direct measures of these variables apparently do not exist, but it is plausible to assume that consumers in stable communities will tend to have more friends and more frequent contact with them. Making friends and contacts will ordinarily require some time after a person has relocated into a new community. Thus our hypothesis is that N is inversely related to measure of population growth and turnover (e.g., the proportion of people who did not live at the same address five years previously). Similarly, social instability variables, such as the proportion of households headed by women, may be related to N, since family instability may reduce the family's social contacts with nonfamily members.

Consequently, to compare the two models, we estimate a price equation for the increasing monopoly model using the proxy variables described above, estimate a price equation for the target income model using sellers per capita as the critical variable, and compare the results in terms of significance and explanatory power. There are, of course, other demand and supply influences on price. Variables to measure their influences are introduced in the next section.

SPECIFICATION OF THE TWO MODELS

Up to this point, the discussion has been limited to consideration of the pricing behavior of given numbers of sellers under the competing hypotheses of the increasing monopoly and target income theories. To ensure that the price equations of each of the models are econometrically identified and to choose the appropriate exogenous variables and estimation techniques, it is necessary to present complete models in which the hypothesized pricing behavior takes place. In particular, we need to be concerned about whether sellers' location and hours of work variables should be treated as endogenous. If they might reasonably be treated as the result of endogenous seller choices, then the appropriate estimation technique for the price equation may need to be based on a simultaneous equations approach. Table 1 provides a list of the variables that will be used in our discussion of the two models.

Table 1. List of Variables

Endogenous variables:		
P	=	seller's unit price
MD	=	number of physicians per square mile
MP	=	number of physicians per capita
H	=	weekly hours available per seller for non-emergency care
Q	=	quantity of services per unit time provided by each seller
Q*	=	quantity of services demanded per capita
D	=	extent of seller's discretionary influence on demand
Exogenous variables:		
MA	=	variables measuring the size of a typical consumer's market area
IS	=	variables measuring the number of information sources available to a typical consumer
W	=	prices of non-physician inputs
A	=	measures of community attractiveness to physicians
Y	=	variables that influence per capita demand for medical care
CL	=	cost-of-living index
PD	=	population density

Increasing Monopoly Theory

Table 2 lists the equations of the increasing monopoly theory. The number of sellers in a community results from sellers' locational decisions. As described by equation IM1, the number of sellers per capita in a community (MP) is a function of the price per unit of

Table 2. Increasing Monopoly Model Specification

Equation Number	Specification
IM1	MP = f_1(P,Q,H,A,W,CL)
IM2	P = f_2(MD,MP,IS,MA,W,Y,CL)
IM3	H = f_3(MD,MP,IS,MA,W,Y,CL)
IM4	Q* = f_4(P,H,Y,CL)
IM5	Q ≡ Q*/MP
IM6	MD ≡ MP/PD
Exogenous variables	A,W,CL,IS,MA,W,PD
Endogenous variables	MP,MD,P,H,Q,Q*

services (P) and the average quantity per unit of time that each physician can sell at that price (Q). Together, these two variables determine each provider's gross income. His net money income per hour then depends on the number of hours he is available to serve non-emergency patients (H), and the prices of inputs he buys, measured by a wage index (W), and a cost-of-living index (CL). Finally, and very important, location depends on real income which will vary with the attractiveness (A) of the community as a place to live. Thus one's locational decision depends on money income potential, which is described by the variables P, Q, H, W, and CL, and on the attractiveness of the location. Attractiveness, in turn, depends on the community's geographical setting, its climate, its air quality, and the quality of its public services, especially primary and secondary education.

Equations IM2 and IM2 describe the seller's decisions on the competitive variables that he controls: the price (P) that he charges and the number of hours (H) that he is available to patients.[2] Note that H measures not only the time the seller actually works, but also idle time he spends waiting for patients, either in the officer or "on call." The increasing monopoly theory hypothesizes that the seller chooses these control variables by maximizing his real income, given local cost conditions and given the firm-level demand curve he faces. The position and shape of that demand curve depend in turn on local competitive conditions, described by seller density per unit area (MD), variables determining the size of each consumer's market area (MA), variables proxying the number of information sources in each market area (IS), and determinants of the overall level of demand in the market area (Y, MP, and CL). Under the assumption of marginal cost (within which the opportunity cost of the physician's foregone leisure is implicitly included) increasing, price will tend to be *negatively* related to physicians per capita (MP), a prediction which is in contrast to that of the target income model, which permits a positive relationship. As above, local marginal costs are measured by input prices (W) and the cost-of-living index (CL).

Equation IM4 describes how consumers respond to these seller decisions. The quantity each buyer demands per unit time is a function of the price charged (P), the hours sellers are available to non-emergency patients, (H), and the aggregate level of consumer demand (Y). Both consumer income and insurance coverage would also be expected to influence the level of demand per capita. Per capita demand is then related to per seller demand by the identity IM5. Finally, equation IM6 is an identity specifying the relationship between MP and MD.

In general form, this system is similar to that suggested by Fuchs and Kramer [3], but with three major differences. First, the market is explicitly treated as one in which individual sellers have monopoly power rather than, as in Fuchs-Kramer, as one in which physicians are price takers. That is, in our model (and generally in reality), each seller can set his price, but cannot sell as much as he wants at that price. Second, variables which describe firm-level demand curves are explicitly included. Finally, differences in input prices are taken into account in explaining location and price.

The Target Income Theory

The specification of the modified target income theory, as described by Sloan and Feldman [12], Evans [2], and Pauly [8], is given in Table 3. This theory differs from orthodox pricing theories in assuming that sellers are only partially constrained by a given consumer demand function. Sellers are assumed to be able to induce buyers to purchase more, from each seller and in total, at a given price. However, this inducement or demand creation, requires some reduction in the accuracy of information provided by the seller, and such reduction has a utility cost to the seller. While there is probably some upper limit to the extent of possible demand creation, within that limit sellers have considerable discretion over the money income they receive. Higher incomes do, however, have a cost in terms of greater manipulation or distortion of information provided (explicitly or implicitly) to patients, and such manipulation has a psychic cost to the provider.

Equation TI1 therefore makes the locational decisions of providers depend on real money income, hours of work (H), community attractiveness (A), *and* the amount of demand manipulation (D) that sellers in a community typically must perform in order to earn such an income. As in the previous model, real net income depends on price (P), quantity sold per week (Q), the hours that the physician is available for non-emergency care (H), and community attractiveness (A).

Equations TI2, TI3, and TI4 describe how the seller maximizes his utility, given the loose market constraints

[2] A more general formulation of the model would allow the provider to choose the length and quality of visit he provides. Nevertheless, since the data used did not permit us to take this into account, specifying a more general and complete model than we specify here would only add complexity, not substance, to our analysis.

that are hypothesized by this theory. A seller's price, the degree to which he influences demand, and the hours he is available for service are functions of the relative supply of sellers (MP), the determinants of per capita demand (Y), and the costs of inputs (W and CL). Finally, equations TI5 and TI6 describe per capita and firm-level consumer demand. The specification is the same as in the increasing monopoly model except that demand from any seller also depends on the seller's level of discretionary activity.

Table 3
Extended Target Income Model Specification

Equation Number	*Specification*
TI1	$MP = f_1(P,Q,H,A,W,CL,D)$
TI2	$P = f_2(MP,W,Y,CL)$
TI3	$H = f_3(MP,W,Y,CL)$
TI4	$D = f_5(MP,W,Y,CL)$
TI5	$Q^* = f_4(P,H,Y,CL,D)$
TI6	$Q \equiv Q^*/MP$
Exogenous variables	A,W,CL,Y
Endogenous variables	MP,P,H,Q,Q*,D

Identifiability and Estimation

With the exception of TI1 and TI5, all equations in both models are identified according to the order conditions of identifiability.[3] Data limitations prevent us from estimating each complete system of equations. In particular, we will only be able to estimate the price equations. In the IM model, both MD and MP, which appear in the price equation, are endogenous. Consequently, the use of ordinary least squares (OLS) gives biased and inconsistent estimates, and a two-stage estimating procedure (2SLS) is therefore to be preferred. Similarly, in the TI price equation, MP is endogenous. In the empirical results which follow, we shall therefore present estimates obtained by using 2SLS. OLS estimates are presented for comparison.

VARIABLE DEFINITIONS AND DATA SOURCES

The data for which we have estimated the price equations of both the increasing monopoly model and the target income model are a cross section of 100 of the largest SMSAs in the United States.[4] The independent variable is an index of the fee charged for a "routine office visit" from a primary care physician. This is an appropriate and tractable set of prices to use for testing the two theories because insurance coverage, excepting Medicaid and Medicare, is uncommon. This fee

[3] If the exogenous variables Y, A, and W are each vectors with sufficient numbers of components, then TI1 and TI5 are also identified.

[4] Absence of data on some variables reduced the final sample size slightly, to 92 cities. Seven cities in New England, plus Flint, Michigan, were deleted.

Table 4. Variable Names, Descriptions, and Sources

Endogenous Variables		
P1973	Price index for routine office visit to a primary care physician, 1973	Mathematica, Inc. telephone survey
MDPCM2	Primary care physicians per square mile of urbanized area, 1970	American Medical Association Survey and *County and City Data Book, 1972* (CC72)
MDPCPC	Primary care physician per capita, 1970	AMA survey and CC72
Exogenous Variables		
Consumer Information (in addition to MDPCM2)		
FEMH	Percent of families that have female heads, 1970	CC72
MOVED	Percent of housing units occupied by residents who moved into unit during 1965-1970	CC72
PUBTR	Percent of workforce using public transport to reach work, 1970	CC72
Demand Determinants		
FAMINC	Median family income, 1970	CC72
AGED	Percent of population 65 or over, 1970	CC72
KIDS	Percent of population under 5, 1970	CC72
BLACKS	Proportion of population that is black, 1970	CC72
SCHOOL	Median years of schooling for those 25 and over, 1970	CC72
PAFDC	Percent of population receiving aid to families with dependent children	CC72
Proportion General Practitioners		
GPMDPC	Proportion of primary care physicians who are general practitioners, 1969	AMA survey
Attractiveness of SMSA		
ATT	Hotel expenditures per capita, 1967	CC72
PROFMAN	Percent of workforce who are professionals or managers, 1970	CC72
LOCGOV	Per capita direct, general expenditures by local government, 1967	CC72
Other Variables		
W	Average industrial wage rate for production workers, 1967	CC72
POP	Total population, 1970	CC72
ADJDEN	Population per square mile within the urbanized area of the SMSA, 1970	CC72
CL 73	Cost of living, 1973	Bureau of Labor Statistics

information was collected by telephone survey in November and December 1973. Data regarding the number of primary care physicians per square mile and the number of physicians per capita in each SMSA were obtained from an American Medical Association survey of physicians done in December 1969. The other variables used, which describe each SMSA's economic, social and demographic characteristics, were obtained from the *County and City Data Book, 1972.* With a few exceptions, all data used from it are based on the April 1970 United States Census of Population. Details of the variables and their sources are as follows. Table 4 summarizes these details; Table 5 lists the means and standard deviations of all variables.

Table 5.
Means and Standard Deviations of Variables

Variable[a]	*Mean*	*Standard Deviation*
P1978	8.45	1.21
MDPCM2	1.48	.79
MDPCPC	.0004	.0001
FEMH	11.	1.9
MOVED	53.	7.5
PUBTR	7.0	7.1
FAMINC	1,002.	122.
AGED	9.1	2.4
KIDS	8.5	.71
BLACKS	.12	.087
SCHOOL	12.	.4
GPMDPC	.50	.11
ATT	3.2	2.2
PROFMAN	24.	3.4
LOCGOV	243.	70.
W	3.0	.47
POP	1,176,870.	1,626,960.
ADJDEN	3,203.	1,233.
PAFDC	5.0	1.8

[a]Some variables are scaled; e.g., FAMINC.

Price of Primary Care Physicians' Services

Data on primary care physicians' fees were obtained from Wooldridge's report [14, Table 2A] on a telephone survey of physicians' offices in 100 large SMSAs that Mathematica, Inc., conducted in December and November 1973. Wooldridge in her report lists for 100 large SMSAs weighted averages of the prices that surveyed general practitioners, internists, and pediatricians reported themselves as charging for a routine office visit in November and December 1973. The weights used in constructing the average for each SMSA were the relative proportions of general practitioners, internists, and pediatricians actually occurring in that SMSA.

Primary Care Physicians Per Capita and Physicians Per Square Mile

The total number of primary care physicians in each metropolitan area as of December 31, 1969, was obtained from American Medical Association data [4, Table 7]. Specifically, those data listed by metropolitan area both the number of office-based general practitioners and the number of office-based medical specialists; the sum of these two figures provides a measure of the number of office-based primary care physicians. This sum, of course, includes a certain number of physicians, such as psychiatrists, who are office-based medical specialists but who do not deliver primary care.

Offsetting this bias, to some degree, are those surgical specialists and other hospital-based physicians who deliver some primary care.

The number of primary care physicians per capita (MDPCPC) was obtained by dividing the estimates of the total number of primary care physicians by the SMSA's total population (item 3, Table 3, *County and City Data Book, 1972;* abbreviated [13, T3, I3] henceforth). The number of primary care physicians (MDPCM2) per square mile within the urbanized parts of each SMSA was calculated by multiplying the primary care physicians per capita (MDPCPC) by the population density of the SMSA within its urbanized subareas (ADJDEN: 13, T4, I204].[5] The population density within the urbanized subarea of each SMSA was used instead of the population density within the entire SMSA because the geographical boundaries of an SMSA often include large amounts of rural land, since SMSAs are defined in terms of counties. Urbanized areas, however, are defined in terms of a certain thickness of settlement [13]. Consequently, the urbanized population density gives a much better estimate of how heavily population within an SMSA is concentrated than that provided by the overall SMSA population density.[6] In addition, adjusted population density within the urbanized area (ADJDEN) was entered explicitly, since this variable may proxy differences in travel time or cost and hence the size of the market area.

Consumer Information Proxies

According to the increasing monopoly model developed above, price is inversely related to the degree

[5]In some cases, there is not a one-to-one correspondence between SMSAs and urbanized areas. For example, the Los Angeles-Long Beach urbanized area is composed of parts of the Los Angeles-Long Beach SMSA, the Anaheim-Santa Ana-Garden Grove SMSA, and the San Bernardino-Riverside-Ontario SMSA. Therefore, the adjusted density (ADJDEN) value used for the Anaheim-Santa Ana-Garden Grove SMSA was the density value listed for the Los Angeles-Long Beach urbanized area because no separate urbanized area is defined for that SMSA. A separate urbanized area (the San Bernardino-Riverside urbanized area) is defined within the San Bernardino-Riverside-Ontario SMSA; therefore, for that SMSA the density value for that latter urbanized area was used.

[6]Some SMSAs included within the data that we used have essentially 100 percent of their populations within areas classified as urban, while others have only 50 percent of their populations within areas classified as urban. Consequently, MDPCM2, as defined above, may not be a good measure of population density within those SMSAs where the percent of urban population is relatively low. To check the importance of this: the increasing monopoly model was estimated using only those SMSAs having populations that area at least 85 percent urban. The results using this subset of the data were essentially identical to the results reported on Table 6, except that the t values were smaller because of the substantially smaller sample size (N = 56 instead of N = 98).

of consumer information and directly related to the degree of consumer ignorance. Proxies for the degree of consumer ignorance are physician density (MDPCM2) as defined above, the percent of families that have a female head [FEMH: 13, T3, 151], and the percent of occupied housing units whose occupants had moved into the unit during the preceding five years [MOVED: 13, T3, 193]. Additionally, the proportion of workers who used public transit to reach their jobs during the week preceding the census [PUBTR: 13, T3, 148] may be an indication of consumer mobility within the SMSA, i.e., a high proportion of workers using public transit may indicate that travel speed in the SMSA is low. Consequently it may be an imperfect, inverse measure of the real extent of the area in which the typical consumer shops for a provider. In addition, PUBTR may also embody information about the SMSA's geographical configuration. The rationale for each of these variables was discussed above, under "Towards Empirical Estimation."

Table 6. Price Equation Estimates: Increasing Monopoly Model

Variable	Regression Coefficients (t statistics in parentheses)				
Estimation Procedure	OLS	2SLS	2SLS — Logarithms of Monetary Variables		
MDPCM2	2.61 (3.71)	4.24 (2.58)	.552 (2.72)	.753 (1.10)	1.95 (0.60)
ADJDEN	-.0009 (-2.54)	-.0016 (-2.11)	-.0021 (-2.24)	-.0003 (-1.05)	-.0008 (-0.60)
FEMH	.22 (2.42)	.23 (2.38)	.032 (2.50)	.087 (1.08)	.22 (0.60)
MOVED	.086 (7.19)	.093 (6.58)	.011 (6.32)	.024 (0.97)	.044 (0.55)
PUBTR	-.040 (-2.36)	-.050 (-2.38)	-.006 (-2.26)	.004 (0.37)	.019 (0.45)
FAMINC or ln FAMINC	.00273 (2.17)	.0035 (2.24)	.49 (2.41)	.87 (0.73)	3.94 (0.52)
AGED	.11 (2.11)	.10 (1.85)	.009 (1.29)	-.020 (-0.30)	-.005 (-0.32)
GPMDPC	-1.8 (-2.20)	-2.1 (-2.34)	-.28 (2.39)	.040 (0.09)	-.081 (-0.07)
W or ln W	.61 (2.68)	.69 (2.68)	.26 (2.66)	.57 (1.17)	.87 (0.64)
PAFDC	-.012 (-0.22)	.017 (0.26)	.004 (0.44)	-.011 (-0.34)	-.037 (-0.38)
BLACKS	-.483 (-0.35)	-1.66 (-.91)	-.315 (-1.40)	-1.34 (-0.50)	-6.37 (-0.46)
SCHOOL	-.20 (-0.71)	-.38 (1.12)	-.04 (0.94)	-.11 (0.44)	-.39 (-0.42)
MDPCPC	-9320 (-3.28)	-17270 (-2.31)	-2533 (2.74)	-4803 (-1.05)	-1315.3 (-0.63)
CONSTANT	6.51 (1.49)	11.9 (1.82)	-.154 (-0.11)	-1.47 (-0.27)	-16.8 (0.46)
CL 73					-3.92 (-0.51)
N	92	92	92	29	29
$\bar{R}^2$(Adjusted)	.697				
DEP VAR	P1973	P1973	ln 1973	ln P1973	ln P1973
Excluded, exogenous variables used in first stage of 2SLS		SCHOOL ATT PROFMAN LOCGOV POP	SCHOOL ATT PROFMAN LOCGOV POP	SCHOOL ATT PROFMAN LOCGOV POP	SCHOOL ATT PROFMAN LOCGOV POP

Industry Demand Determinants

Both the target income model and the increasing monopoly model predict that the overall level of demand for medical care in the SMSA is likely to affect the pricing decisions of providers. The five variables that are included in the analysis because they are determinants of the aggregate demand for medical care within each SMSA are: median family income [FAMINC: 13, T3, I58], percent of population over 65 [AGED: 13, T3, I14], percent of population under 5 [KIDS: 13, T3, I12], proportion of population that is black [BLACKS], percent of persons receiving benefits from the aid to dependent children program (and eligible for Medicaid) [PAFDC: 13, T3, I72], and the median years of schooling completed by persons 25 years old and over [SCHOOL: 13, T3, I24].[7] These variables, especially KIDS and SCHOOL, may also be determinants of the level of consumer information, but *a priori*, their primary effect would appear to be on aggregate demand.

Proportion of General Practitioners

The price data used, as described above, are a weighted average of general practitioner fees, interest fees, and pediatrician fees. Since general practitioners tend to charge lower fees than primary care medical specialists it is important to control for the proportion of primary care physicians who are general practitioners. This variable (GPMDPC) was calculated by dividing the number of office-based general practitioners [5, Table 7] by the total number of primary care physicians as defined above.

Other Variables

Input prices may affect prices primary care physicians charge. The one measure of input prices used was the average hourly industrial wage in the SMSA (W).[8] Providers' location decisions are affected by the SMSA's attractiveness as a place to live and work. Fuchs [4] has suggested that hotel receipts per capita (ATT) within the SMSA is a reasonable indicator of attractiveness.[9] The idea is that people visit relatively attractive cities (and spend money on hotel accommodations) more often than relatively unattractive cities. Other variables that influence the attractiveness of a city are the percent of the labor force that are professionals or managers [PROFMAN: 13, T3, I45] and the per capita, direct general expenditures by local government in 1967 [LOCGOV: 14, T3, I109].[10]

Missing Data

Reliable measures of the comparative cost of living in the full sample of SMSAs do not exist. There are cost-of-living indexes for 29 of the 92 SMSAs in our sample; however, an analysis of these data, to be discussed below, suggests that omission of this variable does not lead to serious specification error.

EMPIRICAL RESULTS

Specification of the Regression Equations

As discussed above under "Specification of the Two Models," the measures MDPCPC and MDPCM2 of physician stock are likely to be endogenous. Therefore the price equations for both models are estimated using 2SLS. OLS estimates are also included in Tables 5 and 6 for comparison.

The variable GPMDPC, the proportion of primary care physicians, is classified as an exogenous variable in all the 2SLS regressions. In a more complete model it would be endogenous because the comparative levels of general practitioner fees versus medical specialists' fees generally might affect the relative numbers of general practitioners and medical specialists who practice in a SMSA. In this model, however, GPMDPC is correctly considered exogenous because the price index that is used as the dependent variable is a weighted average of general practitioner and primary care medical specialist fees. Consequently, high values for the index in a particular SMSA do not convey information about the differential attractiveness of that SMSA to general practitioners and to specialists. A high value of P only means, *ceteris paribus*, that on average the SMSA is attractive to primary care physicians as a group.

Increasing Monopoly Theory

Table 6 presents the regression results for the price equation of the increasing monopoly model. These results are consistent with the increasing monopoly model. Specifically, the variables that relate to consumer information levels (MDPCM2, FEMH, MOVED, PUBTR and ADJDEN) all have the expected signs and are significant at the 5 percent level. The negative sign on ADJDEN is consistent with the view that consumers who are prevented by high travel costs from using a large number of different physicians will have better information on the smaller number of physicians they do use. This "small" number is apparently large enough, and there is apparently enough of a linkage between market areas in a metropolitan area, to avoid oligopoly effects on price. Additionally, the variables that relate to aggregate industry demand

[7] BLACKS was calculated by dividing the number of Negroes in the SMSA [13, T3, I10] by the total population [13, T3, I3].

[8] W was calculated by dividing total wages for production workers in 1967 [13, T3, I128] by the total man-hours for production workers in 1967 [13, T3, I127]. These data are based on the *1967 Census of Manufacturers*.

[9] ATT was calculated by (a) multiplying total receipts for selected service establishments [13, T3, I151] by the percentage of those receipts that were collected by hotels, motels, and camps [13, T3, I154] and (b) then dividing by total population [13, T3, I3]. These figures are based on the *1967 Census of Business*.

[10] The primary source for LOCGOV was the *1967 Census of Governments*.

all have easily interpretable coefficients. For example, the percent of the population under 5 years of age has a significant, negative coefficient while the percent of the population 65 and over has a significant positive coefficient. The negative sign of the former coefficient may reflect that pediatric office care tends to be uninsured and the mothers of young children tend to be unusually aggressive and social consumers of care for their children. The positive sign of the latter coefficient may reflect that care for the aged is largely insured through Medicare and that the aged tend to be socially isolated. The coefficient on PAFDC is not significant. The OLS equation explained 70 percent (adjusted) of the variation of the dependent variable P1973. The coefficient on the physician-population ratio is negative and significant. High numbers of physicians per capita do depress price (by lowering the opportunity cost of physician time), just as the increasing monopoly theory suggests.

Target Income Model

The results in Table 7 indicate why the target income model is deceptively attractive. The price equation of the TI model does not include any variables not present in the IM model, but it does omit some of those variables. When the information source and market area variables are omitted, physicians per capita (MDPCAC) *is* positively related to price, and is significant in 2SLS regressions, although not in the OLS ones. As suggested by the results for the IM model, however, this positive effect disappears when information and market area variables are added. Moreover, the set of such variables is highly significant, as indicated by the rise in $\overline{R}^2$ from .43 to .70. The F statistic for the set of variables included in IM but not in TI is 15.6, significant at better than 99 percent for 5 and 77 degrees of freedom.

Table 7. Price Equation Estimates — Target Income Model

Variable *Estimation Procedure*	2SLS	OLS
MDPCPC	8054	2547
	(2.46)	(1.59)
GPMDPC	-1.8	-2.4
	(-1.55)	(-2.23)
W	.17	.034
	(0.52)	(-0.12)
FAMINC	.0001	.0021
	(0.08)	(1.67)
AGED	.087	.049
	(1.22)	(0.76)
KIDS	.049	-.40
	(-0.17)	(1.75)
BLACKS	4.6	3.1
	(2.52)	(1.98)
SCHOOL	.53	.59
	(1.67)	(1.78)
PAFDC	.035	.121
	(0.43)	(1.92)
Excluded exogenous variables used in first stage of 2SLS	POPGR POP ATT LOCGOV PROFMAN	
N	92	92
$\overline{R}^2$		.431

COST-OF-LIVING ADJUSTMENTS

Areas differ not only in the prices that prevail for medical care, but also in the prices of all other goods and services. In principle, the price discussed in the theoretical model is the *relative* price of physicians' services, and so some adjustment should be made for possibly offsetting differences in the prices of other goods. In addition, the general price level affects the level of real income of consumers and physicians, and these effects should be taken into account.

Unfortunately, there is no cost-of-living index or similar index of general price levels for all of the 92 cities in our data. Such an index may be effectively provided to some extent by the wage rate variable, since levels of area wage rates and of area prices tend to be positively related. Indeed, if workers reach a movement equilibrium in which the real value of money income per unit of labor is equalized across areas, the wage rate will be proportional to the cost of living. Even in the absence of such an equilibrium, the wage index may account for much of the variation in living costs across areas.

A cost-of-living index is published for 29 of the cities in the sample (CL 1973, high income family budget in [1]), and so the effect of its inclusion in the IM model could be determined for that (small) subsample. A flexible method for including CL is to express the index and all monetary variables (independent and dependent) in logarithms. This specification permits (but does not require) the CL index to have an effect on real prices which differs from that of the monetary variables.[11] As might be expected with a small sample, when the IM equation is estimated with the data from 29 cities, the significance level of individual varibles tends to be low. However, as can be seen in a comparison of columns 4 and 5 of Table 6, the coefficients tend to have the same signs; only 2 of 9 significant coefficients (those on PUBTR and GPMDPC) change in sign. If this

[11] If the cost-of-living variable is an equally appropriate deflator for all monetary variables, the estimating equation can be written (in the monetary variable wages W and income Y only) as

$$\frac{P}{CL} = \left(\frac{W}{CL}\right)^{\alpha_1}\left(\frac{Y}{CL}\right)^{\alpha_2}$$

Taking logarithms, one obtains:

$$\ln P - \ln CL = \alpha_1 \ln W + \alpha_2 \ln Y - (\alpha_1 + \alpha_2) \ln CL$$

This yields:

$$\ln P = \alpha_1 \ln W + \alpha_2 \ln Y - (\alpha_1 + \alpha_2 - 1) \ln CL$$

If CL is an appropriate deflator for all variables, then the coefficient on CL should approximately equal

$$-(\alpha_1 + \alpha_2 - 1).$$

subsample is thought to be similar to the full sample, the results in column 6 of Table 6 suggest that adding a cost-of-living index (CL 1973) does not appreciably affect the qualitative results. Only one variable (GPMDPC) switches in sign, and coefficients of the other variables tend to increase. The cost-of-living variable itself is not significant.[12]

CONCLUSION

The empirical results in the study should, of course, be regarded as preliminary. Better measures of consumer information on fee levels, and of input prices are desirable. The data on the cost of living across SMSAs are quite incomplete. Nevertheless, the conformity of the results with the increasing monopoly model is striking. Not only is the measure of physician stock suggested by that model far more useful in explaining price than the physician-population ratio, but other variables that could only have been suggested for inclusion by the increasing monopoly model, such as the percent of families headed by females, are highly significant. Judged on both of these grounds, the increasing monopoly must be regarded as a strong competitor with the target income model (and with the neoclassical competitive model) in explaining price formation for physicians' services. Moreover, it appears possible that the increasing monopoly model could provide an explanation of the pricing behavior of other service industries.

The main implication of these results is that consumers may, in fact, through their market choices, exert substantial influence over the pricing decisions of primary care physicians. In other words, these results suggest that if consumers have access to comparative information about primary care physicians, they may collectively be far from powerless in their dealings with primary care physicians. These results, however, do not give practical advice as to how consumer access to information about competing physicians can be improved.

References

1. Brackett J: Urban family budgets updated to autumn, 1973. *Monthly Labor Review* 97:57-62, 1974
2. Evans R: Supplier-induced demand. In: *The Economics of Health and Medical Care* (Perlman M, ed.). London, Macmillan, 1974, pp 162-173
3. Fuchs VR, Kramer MJ: *Determinants of Expenditures for Physicians' Services in the United States, 1948-68.* New York, National Bureau of Economic Research, 1972
4. Fuchs VR: The supply of surgeons and the demand for surgical operations. *Journal of Human Resources*, in press
5. Haug JN, Roback GA: *Distribution of Physicians, Hospitals, and Hospital Beds in the U.S., 1969: Volume II, Metropolitan Areas.* Chicago, Department of Survey Research, American Medical Association, 1970
6. Huang LF, Koropecky O: *The Effects of the Medicare Method of Reimbursement on Physicians' Fees and Beneficiaries' Utilization.* Washington, D.C., Robert R. Nathan Associates, 1973
7. Newhouse JE: A model of physician pricing. *Southern Economic Journal* 37:174-183, 1970
8. Pauly MV: *The doctor and his workshops*, in press
9. Reinhardt UE: Comments on Sloan and Feldman. In: *Competition in the Health Care Sector: Past, Present and Future* (Greenberg W, ed.), Washington, D.C., Federal Trade Commission, 1978, pp 156-190
10. Satterthwaite MA: The effect of increased supply on equilibrium price: a theory for the strange case of physicians' services. Northwestern University Center for Health Services and Policy Research, November 1977
11. Sloan FA: Physician fee inflation: evidence for the late 1960's. In: *The Role of Health Insurance in the Health Services Sector* (Rosett RR, ed.). New York, National Bureau of Economic Research, 1976, pp 321-354
12. Sloan FA, Feldman R: Monopolistic elements in the market for physicians' services. In: *Competition in the Health Care Sector: Past, Present and Future* (Greenberg W, ed.). Washington, D.C., Federal Trade Commission, 1978, pp 57-131
13. U.S. Bureau of the Census: *County and City Data Book, 1972.* Washington, D.C., U.S. Government Printing Office, 1973
14. Wooldridge J: *The Price of Medical Care as Reflected in a Survey of the 100 Largest Standard Metropolitan Statistical Areas.* Policy Analysis Series No. 10, Mathematica, Inc., November 1975

[12] Adding the coefficients on W and Y and subtracting 1 yields 3.8, approximately equivalent in value and opposite in sign to the 3.9 coefficient on CL. As discussed in note 11, this result indicates that CL is functioning as a simple cost-of-living deflator for all variables.

The Aggregate Supplies and Demands of Physician and Dental Services

Jesse S. Hixson, Ph.D.*
Nina Mocniak, M.P.H.*
Health Resources Administration

ABSTRACT □ A two-equation model of the supply and demand for dental services and a four-equation model of the joint determination of the quantities and prices of physician and hospital services are investigated with aggregate time-series data, 1950-1970 and 1949-1975, respectively. Instrumental variables with the Hildreth-Lu autoregressive technique are employed; separate parameters for the income and price elasticities of the insured and uninsured populations are estimated.

Conventional economic behavior is evident on both the demand and supply sides of the markets. In particular, the estimated price elasticities of supply are 0.71 and 1.54 for physician and dental services, respectively. The estimated elasticity of supply is unitary with respect to the aggregate stocks of both physicians and dentists.

INTRODUCTION

In this paper, we investigate the aggregate supply and demand functions of physician and dentist services. The properties of these functions have both short-run implications for responses to increases in the demands for services which might be precipitated by national health insurance, and long-run implications for the consequences of health manpower policy. Recently, the interest in the properties of these functions has intensified with the renewed debates over national health insurance and manpower policy, and with the recent emergence and popular acceptance of the "target income" hypothesis respecting physician and dentist supply behavior and the role of market forces in allocating health care resources. We hope that the empirical research presented here will help to resolve some of the issues which revolve around the properties of these functions.

The empirical work reported here departs from recent approaches to the issues in question in that it is based on aggregate time series rather than cross-section data. One of the first studies of the physician market was undertaken by Feldstein [1] with aggregate time-series data spanning the period 1948-66. Feldstein's conclusions from his study have had a profound effect on the thinking of many economists. His paper is often cited as evidence that traditional economic models are not applicable to the markets for health services. Perhaps because of Feldstein's peculiar empirical results the time-series data have tended to be ignored in favor of cross-section observations obtained from survey or census data. As will be seen below, our empirical results — judged from the standard statistical and economic-theoretic viewpoints — are quite good, suggesting that perhaps these time-series data can be further exploited in modeling the markets for health services. Furthermore, our results essentially refute Feldstein's conclusion that orthodox economic models do not characterize the markets for health services.

The remainder of the paper consists of three parts. In the next section we present the formal models from which the estimating equations are derived. In the subsequent section, we present the estimates of the physician and dentist supply and demand functions, respectively. The results are summarized and conclusions are presented in the last section.

*The authors, who are employed by the Bureau of Health Manpower, U.S.D.H.E.W., wrote this paper in their private capacity; no endorsement by U.S.D.H.E.W. is intended or implied.

THE MODELS

The objective of the exercise reported here was to explain the prices and quantities of physician and dentist services using the simplest models of economic theory possible. The quantities are measured by the total real annual expenditures on physician and dental services in the U.S., while the prices are measured by the respective fee components of the Consumer Price Index relative to the overall value of the Consumer Price Index. The basic modeling assumption was that these observed magnitudes were generated by the equilibration of the aggregate demand and aggregate supply of physician and dentist services.

The models treat the aggregate demand for each service as generated by two distinct populations: the "uninsured," who pay the full market price for services; and the "insured" who pay a "net" price which is less than the full market price paid by the uninsured. The concept of the models is illustrated in Figure 1, which depicts the equilibrium condition. In Figure 1, the aggregate demand function D is the sum of the demand of the uninsured, d_u, and the insured, d_i.

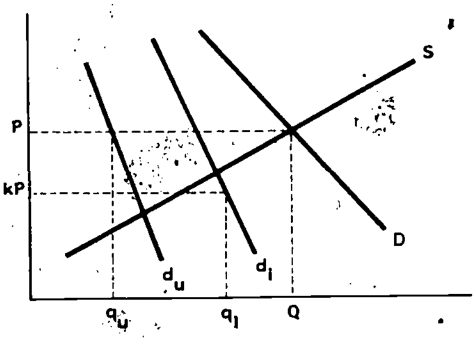


Figure 1

The demand curve of the insured can only be plotted for a specified value of the proportion of the full price paid by the insured, denoted by k; therefore, the position of the aggregate demand curve D is a function of k. For a given supply curve S, the equilibrium full price P and the proportion k of the full price paid by the insured must be such that the sum of the quantity demanded, q_u, by the uninsured and that by the insured, q_i, equals Q, the quantity supplied at the full price P.

One notes that in this model, a coordinate (P,Q) lies on both the aggregate demand function and the aggregate supply function. This is in contrast to models proposed by others, e.g., [1] and [2], that do not explicitly distinguish between the separate demands of the insured and uninsured but treat aggregate demand as an average of the demand of both groups at some "average" aggregate price. In these latter models, the coordinate (P,Q) lies only on the supply curve; for the demand curve to be identified, an unobservable aggregate "average" price must somehow be estimated. The estimation of such an average demand price has been a major stumbling block in previous attempts to estimate the parameters of demand functions for physician and hospital services. Moreover, the confounding of the behavior of the insured and the uninsured in such models has had serious consequences for properties of the empirical estimates of the parameters of the demand functions. In the estimation of the model analyzed in this paper, we have avoided these problems by separating the demands of the insured and uninsured.

The behavioral equations of the models are specified in per capita terms, which restricts the functional form of the aggregate demand and supply equations and enhances the efficiency of the estimates of the parameters. This enhancement derives from the correlations between the variables in the estimating equations being significantly reduced in aggregating the per capita equations to derive the aggregate equations.

The Demand and Supply of Physician Services

Since physician and hospital services are often consumed jointly, the model of physician services includes equations for the demand and supply of hospital services. The behavioral and aggregated equations of the physician services model are as follows.

The average annual quantities of physician and hospital services demanded by each uninsured person are given, respectively, by

$$q_{ut} = a_0 + a_1 p_t + a_2 y_{ut} + a_3 r_t \qquad (1)$$

$$z_{ut} = b_0 + b_1 p_t + b_2 y_{ut} + b_3 r_t \qquad (2)$$

where p_t and r_t are the prices of the respective services and y_{ut} is the per capita income of the uninsured in year t. The average annual quantities of physician and hospital services demanded by each of the insured are given, respectively, by

$$q_{it} = a_4 + a_5 k_t p_t + a_6 y_{it} + a_7 l_t r_t \qquad (3)$$

$$z_{it} = b_4 + b_5 k_t p_t + b_6 y_{it} + b_7 l_t r_t \qquad (4)$$

where k_t and l_t are the proportions of the respective prices paid by the insured and y_{it} is their per capita income. The demands for each type of service within each population group are interrelated through the prices of both services appearing in each demand equation.

The supply of physician services per physician is given by

$$q_{st} = \gamma_0 + \gamma_1 P_t + \gamma_2 t \qquad (5)$$

where t is included as an argument to determin the current state of technology. The supply of hospital services per hospital bed is given by

$$z_{st} = \lambda_0 + \lambda_1 r_t + \lambda_2 t. \qquad (6)$$

The aggregate demand and supply functions investigated empirically are derived from equations (1)-(6) by multiplying each by the sizes of the relevant populations, and adding the aggregate demand functions of the insured and uninsured populations. The national aggregate demand and supply system thus derived is:

$$Q_{dt} = a_0 N_t(1\text{-}I_t) + a_1 p_t N(1\text{-}I_t) + a_2 Y_{ut} + a_3 r_t N_t(1\text{-}I_t) + a_4 N_t I_t + a_5 k_t p_t N_t I_t + a_6 Y_{it} + a_7 l_t r_t N_t I_t \tag{7}$$

$$Q_{st} = \gamma_0 D_t + \gamma_1 p_t D_t + \gamma_2 t D_t \tag{8}$$

$$Z_{dt} = b_0 N_t(1\text{-}I_t) + b_1 p_t N(1\text{-}I_t) + b_2 Y_{ut} + b_3 r_t N_t(1\text{-}I_t) + b_4 N_t I_t + b_5 k_t p_t N_t I_t + b_6 Y_{it} + b_7 l_t r_t N_t I_t \tag{9}$$

$$Z_{st} = \lambda_0 H_t + \lambda_1 r_t H_t + \lambda_1 t H_t \tag{10}$$

$$Q_{dt} = Q_{st}; Z_{dt} = Z_{st} \tag{11}$$

In the system (7)-(11) Q_t and Z_t are the total quantities of physician and hospital services, respectively; N_t and I_t are, respectively, the size and proportion insured of the national population; Y_{ut} and Y_{it} are the respective incomes of the uninsured and insured populations; and D_t and H_t are the total number of doctors and hospital beds, respectively, in year t. The measure of each variable in system (7)-(11) is as follows:

Q_t,Z_t: The quantities of physician and hospital services were measured by total annual expenditures deflated by the appropriate component of the Consumer Price Index.

p_t,r_t: The respective prices of physician and hospital services are measured by the respective fee components of the Consumer Price Index relative to the overall value of the Consumer Price Index.

N_t: The measure is the total U.S. population.

I_t: Two alternative measures of the proportion of the population "insured" were available: the proportion having hospitalization coverage and the proportion having surgical coverage. Because of the interdependence of the demands for hospital and physician services, neither measure fully encompasses the effect of the "insurance" on the demands for services. Therefore, the choice of an insurance variable was made on empirical grounds. The surgical coverage variable gave the best results in terms of statistical performance, so it is used throughout the estimation procedure as the measure of I_t.

k_t, l_t: No series on the proportions of the respective prices of physician and hospital services paid by the insured are available. Therefore, proxy variables for the true k_t and l_t had to be created. While the respective gross prices, p_t and r_t, have been systematically rising, it was inferred from the ratios of direct payments paid by individuals to total expenditures for physician and hospital services that $k_t p_t$ and $l_t r_t$ have been systematically declining through time. Therefore, various power functions defining k_t and l_t as functions of reciprocals of powers of p_t and r_t were used as proxies for the unobserved k_t and l_t.

Y_{it}, Y_{ut}: The measures are the total personal income of the insured and uninsured, respectively. With the exception of three sample years of the Health Interview Survey (HIS), the incomes of the insured and uninsured populations have not been measured. It was therefore necessary to derive estimated series for Y_{it} and Y_{ut} from three HIS estimates of family income of the insured and uninsured in the years 1963, 1968, and 1974. Regressions of the estimated mean family incomes of the two groups onto various powers of total personal income were used to generate estimated series of mean family incomes. These estimated series were divided by the series of average family size, to estimate per capita income, and the resulting series were multiplied in turn by series of $N_t I_t$ and $N_t(1\text{-}I_t)$ to obtain Y_{it} and Y_{ut}, respectively.

D_t: The number of doctors is measured as the total number of physicians alive in the U.S. Some researchers have used the number of active physicians in similar contexts, but the number of active physicians, being a function of the aggregate level of economic activity and the state of demand for physicians' services, is an endogenous variable.

H_t: The number of hospital beds is measured as the total number of hospital beds in the U.S.

The Demand and Supply of Dental Services

Since the market for dental care is not intimately connected to other health care markets as the market for physician services is, the dental service model considers only the demand and supply of dental services. Another feature of the dental sector that affords further simplification of the model is the relative absence of "insurance." By 1970, only 6 percent of the population was participating in a dental prepayment plan. Consequently, we are able to abstract from differences in the income distributions between the insured and uninsured populations, and to assume that the subsidization of a small proportion of demand shifts the aggregate demand function by a constant amount which is proportional to the proportion of the population insured. These assumptions are reflected in the specification of system (12)-(15):

$$q_{ut} = a_o + a_1 p_t + a_2 y_{ut} \tag{12}$$

$$q_{it} = a_3 \tag{13}$$

$$q_{st} = b_o + b_1 p_t + b_2 t \tag{14}$$

$$Q_{dt} = N_{ut} q_{ut} + N_{it} q_{it}$$
$$Q_{st} = D_t q_{st}$$
$$Q_{dt} = Q_{st} \tag{15}$$

where:

Q_{dt} and Q_{st} are the quantity of dental services demanded and supplied, respectively, measured by the total real annual expenditures on dental services in the U.S.;

q_{st} is the real quantity of dental services supplied per dentist in year t;

D_t is the number of dentists in year t;

q_{ut} and q_{it} are the real per capita quantities demanded by the uninsured and insured population, respectively, in year t;

N_{ut} and N_{It} are the number of people, respectively, in the uninsured and insured population in year t;

p_t is the average price in year t of a unit of dental service, measured by the dental component of the Consumer Price Index relative to the general price level;

y_{ut} is real annual personal income per capita of the uninsured in year t.

The definitions of Q_{dt} and Q_{st} are employed to derive the aggregate demand and supply functions (16) and (17) on the assumption that total personal income is distributed among the insured and uninsured according to their proportion of the total population:

$$Q_{dt} = a_0N_t(1\text{-}I_t) + a_1p_tN_t(1\text{-}I_t) + a_2Y_t(1\text{-}I_t) + a_3N_tI_t \qquad (16)$$

$$Q_{st} = b_0D_t + b_1p_tD_t + b_2tD_t \qquad (17)$$

where N_t is the size of the population in year t, I_t is the proportion of the population covered by dental "insurance," and Y_t is total real personal income in year t.

EMPIRICAL RESULTS

The Supply and Demand for Physician Services

Data for estimating the instrumental and proxy variables and parameters of the model were obtained from standard statistical compilations.[1] The estimates of the physician service model are derived from annual observations compiled for the period 1949-1975. All equations were estimated with the Hildreth-Lu autoregressive technique.

Table 1 shows the simple correlations among variables of the physician service model. The correlation among these variables is extremely high, which has profound implications for the precision of the estimates of the paramters, and makes them very sensitive to the choice of the measure of physician and hospital service prices. In addition to the observed values of the price variables, a number of alternative instrumental variables were employed in estimating the parameters of the model. It was found that measures of the variables that reduced the correlations between the regressors tended to improve the precision of the estimates of the parameters (i.e., reduced the variances of the estimates) and also tended to give estimates with the theoretically correct signs. In most cases, we were able to find measures of price that gave very precise estimates of the parameters. In the following discussion of the empirical results, we report the most precise estimates that we found.

The results of estimating the demand functions for physician and hospital services are shown in Table 2. The observed values of the price variables were used in estimating the equations reported in Table 2. The signs of all the statistically significant estimated coefficients are, as one would expect, on the basis of conventional theory. The demand for each service is seen to vary inversely with its price and that of its complement; and demand increase with income. Overall, the estimates are highly significant — 8 out of 12 of the estimated coefficients are significant at the 99 percent level of confidence. The statistical significance achieved with

[1] All the time-series of the variables, with the exception of I_t, Y_{ut} and Y_{it}, were obtained from [6] supplemented by [7]. I_t was obtained from [3] supplemented by unpublished data from the Health Insurance Institute. For the years 1967-75, the number of persons 65 years of age and over having only Medicare hospital insurance, estimated from information provided in [4], were added to the measure of I_t obtained from the Health Insurance Institute. Y_{ut} and Y_{it} were obtained for the years 1963, 1968, and 1974 from [4].

Table 1. Simple Correlations Among Variables — Physician Model

	p	r	N(1-I)	Y_u	NI	Y_i	D	tD	H	tH
p	1.00									
r	0.97	1.00								
N(1-I)	-0.95	-0.95	1.00							
Y_u	-0.90	-0.91	-0.95	1.00						
NI	0.97	0.95	-0.99	-0.97	1.00					
Y_i	0.88	0.96	0.99	-0.88	0.90	1.00				
D	0.97						1.00			
tD	0.97						0.99	1.00		
H		-0.02							1.00	
tH		0.84							0.10	1.00

Table 2. Estimated Coefficients, Demand Equations

Equation	Coefficient: Constant	Price	Time	Annual Compound Rate of Technical Change	Elasticity of Supply with Respect to Physician
Physician Services					0.98
Coefficient	0.20[c]	0.41[a]	0.006[a]	3/4	
Standard Error	0.08	0.28	0.04		
Elasticity		0.41	0.38		
Hospital Services					
Coefficient	-1.46[b]	0.14	0.032[c]		
Standard Error	0.77	0.15	0.012		
Elasticity					

[a] Significant at 99 percent level of confidence (t > 2.47).

Table 3. Empirical Results, Supply Equations

Equation	Coefficient						
	Insured			Uninsured			R^2
	Net Physician Price	Income	Net Hospital Price	Full Physician Price	Income	Full Hospital Price	
Physician Services							
Coefficient	-2.77[a]	-0.24	-0.064[a]	-9.94[a]	22.02[a]	-0.90	0.9920
Standard Error	0.73	0.32	0.28	3.37	7.94	0.92	
Elasticity	-2.17			-3.41	2.12		
Hospital Services							
Coefficient	-1.01[a]	-0.02	-0.04[a]	-4.69[a]	14.17[a]	-1.29	0.9910
Standard Error	0.30	0.15	0.009	1.60	3.86	0.45	
Elasticity	-0.17		-0.15	-1.48	1.23	-0.33	

[a]Significant at the 90 percent level of confidence (t > 1.31)
[b]Significant at the 95 percent level of confidence (t > 1.70)
[c]Significant at the 99 percent level of confidence (t > 2.47)

the time-series data is remarkable, since each demand equation had eight parameters and was estimated with 27 observations, and since there is a high level of correlation among the regressors. Despite the limited number of observations and the high correlation among the variables, the operation of the market mechanisms depicted by the model are clearly making themselves evident in the data.

The only disappointing aspect of the empirical results is the failure of the estimated coefficients of income in the demand of the insured to be significantly different from zero. We have not yet tried to resolve the reason for this result. However, statistical insignificance of insureds' income is persistent throughout all specifications of the demand equations we have estimated.

It is interesting to note the "high" estimates of the elasticities of demand with respect to the prices of physician services, and to the income of the uninsured, i.e., those with an absolute value greater than one. We have not yet investigated problems of specification which might lead to upward bias in the estimated price elasticities. However, it is clear that most estimates of these price elasticities currently cited are derived from cross-section data which contain significant errors-of-measurement leading to bias in estimated coefficients toward zero. Consequently, while some of the elasticities reported here may seem relatively high in relation to those reported by others from studies of cross-section data, the magnitude of upward bias, if any, cannot be deducted through comparisons with estimates derived from cross-section data. Further work needs to be done on isolating the extent of bias in estimates derived from both types of data.

The results of estimating the supply functions for physician and hospital services are shown in Table 3. The hospital service supply equation was estimated with the lagged value of the observed hospital service price serving as an instrument. The hospital supply function is troubled by a high degree of autocorrelation: the value of the first-order coefficient of autocorrelation is 0.95. Modeling the supply of hospital services requires a much more elaborate treatment than we have given it here.

Despite the high level of correlation among the regressors, the estimated coefficients of the physician service supply function are statistically significant at the 90 percent level of confidence or better. The observed values of the price were used to estimate the equation. The results indicate conventional economic behavior on the part of physicians. The elasticity of demand with respect to price computed at the means of the observations indicates that physicians' short-run supply response is price-inelastic. Systematic technological change, as measured by the coefficient of the time variable, has increased the average physician's supply at a compound annual rate of 3/4 percent.[2] The long-run response of the aggregate supply of physicians' services to an increase in the number of physicians has been unitarily elastic.

The contrast in the results reported in Tables 2 and 3 and those reported by Feldstein [1] is striking. Unable to obtain "satisfactory" results after calculating an untold number of regressions, Feldstein concluded that the market for physician services must be in permanent disequilibrium. He interpreted his results as indicating that physicians act in concern to maintain permanent excess demand in order to enhance their discretionary control over the types of cases they treat. He summarized the significance of his paper by suggesting that because his estimates were at variance with the implications of "traditional economic analysis," a reformulation of policy toward physicians' services and prices might be warranted. Based on his unsuccessful attempt to fit alternative models to aggregate time-series

[2]Over the 27-year period, the physician supply function shifted 21 percent due to technological change, which is approximately a 3/4 percent annual compound rate of increse.

Table 4. Simple Correlations Among Variables — Dental Model

	p	N(1-I)	Y_u	NI	D	tD
p	1.00					
N(1-I)	0.87	1.00				
Y_u	0.97	0.87	1.00			
NI	0.70	0.41	0.77	1.00		
D	0.91				1.00	
tD	0.92				0.99	1.00

data, he concluded that "the institutional setting of medical care and the doctors' personal motivations make the conventional economic models inadequate descriptions of physicians' behavior." The empirical results reported above, however, contradict both his conclusions. The economic behavior on the supply side of the market for physicians makes itself apparent in the data, and is entirely consistent with one's expectations drawn from conventional theory. The physician's supply of services is upward-sloping in the price-quantity plane; and the aggregate supply of physicians' services shifts out as the number of physicians increases. For every 1 percent increase in the total stock of physicians the aggregate supply of services increases 0.98 percent. One notes that this conclusion prevails irrespective of whether the physician's underlying supply of labor or work time is "backward-bending." For the measure of output analyzed is that of the physician's firm, which employs the physician's time as only one input and which can consequently expand production with increased use of other inputs even while reducing the input of the physician's own time.[3]

The Supply and Demand of Dental Services

Data for estimating equations (16) and (17) for dental services consisted of time series covering the period 1950-1975.[4] The equations were estimated with the Hildreth-Lu autoregressive technique. Simple correlations among the regressors are shown in Table 4. As in the physician service model, the correlation between the variables is quite high, making the estimates sensitive to the choice of measures of price. In employing alternative instrumental variables, it was found that use of the regression of p onto the raw variables D, Y, N, I, and T as an instrument for p gave the most precise estimates of the demand function parameters, while the regression of p onto Y gave the most efficient estimate of the supply function.

Table 5 shows the estimates of the coefficients and their estimated standard errors, as well as average elasticities computed at the means of the observed variables.

The estimated coefficients of the demand equation have the theoretically correct signs and are statistically significant at the 99 percent level of confidence. The demand of the uninsured for dental services is income-elastic, and is exceedingly price-elastic according to the estimates from the time-series. The extremely high price elasticity of demand reported here is in marked contrast to estimates reported by others from cross-section and survey data. Again, however, we note that many errors-of-measurement and inconsistencies inherent in survey and typically exploited cross-section data bias estimates obtained from them toward zero. Further work is necessary to reconcile the estimates obtained from the alternative types of data.

The estimated coefficients of the supply function have theoretically correct signs and although not as precise as those of the demand function, are statistically significant at the 90 percent level of confidence. The price elasticity of supply calculated at the means of the

[3] For a theoretical exploration of the conditions under which output can increase in response to a change in price despite a decrease in the input of physician's time, see [5].

[4] All the time series of the variables, with the exception of I_i, were obtained from [6] supplemented by [7]; I_i was obtained from [3]. The series of D_t prior to 1958 was revised to achieve consistency in the exclusion of new graduates from current year estimates of the stock of dentists.

Table 5. Estimated Parameters of Supply and Demand Functions for Dental Services

Equation	*Variable*	*Coefficient*	*Standard Error*	*Elasticity*[a]
Demand-uninsured	Constant	4.92	1.55	...
Demand-uninsured	Price	-6.68	2.19	-4.18
Demand-uninsured	Income	2.67	0.59	2.14
Demand-insured	Constant	3.70	0.92	...
Supply	Constant	-4.88	2.57	...
Supply	Price	4.62	2.84	2.88
Supply	Time	0.026	0.019	1.04

[a] For every 1 percent increase in the number of dentists, the aggregate supply of dental services increased by an average of 0.95 percent.

observations is 2.88. Over the period, the aggregate supply of dental services has increased 0.95 percent for each 1 percent increase in the total number of dentists. Finally, technological change accounted for a 54 percent increase in the supply of services over the 26-year period, a compound annual rate of about 1.75 percent.

SUMMARY AND CONCLUSIONS

We have reported the empirical results of our estimating aggregate supply and demand functions of physician and dental services using national aggregate time-series data. The equations estimated were parts of simultaneous-equation systems. First-order autoregressive models were employed to obtain estimates of the parameters of each function. The empirical results indicate conventional economic behavior on the part of both physicians and dentists.

The estimated price coefficients in both supply functions were positive and statistically significant; the estimated price elasticities of supply were 0.41 and 2.88 for physician and dental services, respectively. The implication of these results is that an increase in the aggregate demands for services can call forth supply responses through price-equilibrating mechanisms in the short run. Irrespective of whether or not the individual practitioner's supply of time is backward-bending, the output of the firm is evidently responsive to increases in product price.

With respect to the long run, increases in the stock of practitioners result in increases in the supplies of their services. The estimated elasticity of long-run supply is essentially unitary with respect to the aggregate stocks of both physicians and dentists. Consequently, health manpower policy designed to stimulate the production of physicians and dentists can be expected to lead to increases in the supplies of services as well as to exert significant downward influences on service prices by virtue of the high price elasticities of demand for both medical and dental services.

As on the supply sides of the markets, behavior on the demand sides can be given traditional economic interpretations. The demand for dental services is elastic with respect to both income and price. The demand for physician services is generally dependent on price and income, and on the prices of hospital services. Even the "insured" are seen to be very sensitive to changes in the net prices of physician services. Although less elastic with respect to price and income, the demand for hospital services is sensitive to the price of physician services.

From the empirical results presented in this paper, it is clear that an orthodox price-equilibrating economic model can plainly resolve behavior in the markets for health services. These results clearly refute Feldstein's work [1] on physicians' services, a work that is frequently cited as evidence that orthodox economic models are not applicable to the market for physician services, and by implication, not applicable to the market for dental services either. His reported inability to detect the operation of traditional market forces in time-series data has led many to believe that market forces are either absent or operate in perverse ways in the health care system. The results of the modeling investigation reported in this paper, however, demonstrate that the operation of market forces clearly makes itself apparent in the aggregate time-series data.

References

1. Feldstein MS: The rising price of physicians' services. *The Review of Economics and Statistics* LII:121-133, 1970
2. Fuchs VR, Kramer MJ: *Determinants of Expenditures for Physicians' Services in the United States, 1948-68.* National Bureau of Economic Research, December 1972
3. Health Insurance Institute: *Source Book of Health Insurance Data 1976-77.* New York, 1977
4. National Center for Health Statistics: *Vital and Health Statistics; Data from the National Health Survey: Hospital and Surgical Insurance Coverage.* Series 10, U.S. DHEW
5. Ro KK, Powers JA: A behavior model of physicians: *Korean Economic Journal* 24:37-44, 1976
6. U.S. Bureau of the Census: *Historical Statistics of the United States, Bicentennial Edition, Part 1.* Washington, D.C., 1975
7. U.S. Bureau of the Census, *Statistical Abstracts of the United States* for the successive years 1971, 1972, 1973, 1974, 1975, 1976. Washington, D.C.

The Effect of Local Physician Supply on the Treatment of Hypertension in Quebec*

Philip J. Held, Ph.D.
Larry M. Manheim, Ph.D.
Mathematica Policy Research

ABSTRACT ☐ The primary objective of this analysis is to estimate the impact of the local physician supply on the annual cost and revisit rate for the treatment of essential hypertension. The analysis is based on the medical utilization of a sample of 150,000 beneficiaries in Quebec covering the period from the inception of universal health insurance in 1971 to 1975. The sample, of which approximately 13,000 had been seen for essential hypertension during a one-year period, was stratified by location, age, sex, and family income.

Holding patient age, sex, income, location and a proxy measure of health status constant, we can report the following regarding treatment of hypertension. The effect of the local GP to population ratio is to increase cost. The effect of the local specialists to population ratio is to reduce cost and the revisit rate, although the magnitude of the effect is very small. Holding constant the local supply of physicians, costs are slightly higher, but the revisit rate is lower if the treating physician is a specialist. Costs are rising at approximately 10 percent per year.

These results show that the treatment of hypertension is not independent of the local physician supply. The cost of treatment increases as the number of GPs increases, but decreases as the number of specialists increases. The magnitude of both these effects, however, tends to be small. The change in costs is associated more with the cost per visit rather than the number of visits. Low income persons receive more services per year than do the rest of the population, although these differences are decreasing over time.

INTRODUCTION

"One of the foremost health problems today is hypertension (high blood pressure). It is the most important single risk factor for that class of disease — cardiovascular disease — that kills and cripples more people than any other. Effective treatment is available but infrequently applied."[1]

Ninety-five percent of hypertensives have what is called "essential hypertension," which means nothing more specific than the presence of high arterial blood pressure. Despite the associations it often conjures up, it does not imply any psychological syndrome. Frequently, in fact, there are no outward manifestations, either physical or emotional. Hypertension is not a disease, it is simply a quantitative deviation of blood pressure relative to the norm for a given population.

There are, however, clinical consequences of high blood pressure, which manifest themselves primarily in the brain, the heart, and the kidneys. Brain strokes are likely; coronary artery disease develops at an accelerated rate, manifested in exertional chest pain (angina pectoris), and heart attack (myocardial

*The research reported herein was supported in part by the U.S. Department of Health, Education, and Welfare (National Center for Health Services Research), under contract numbers HRA-230-75-0166 and HRA-230-75-0167. The opinions and conclusions expressed herein are solely those of the authors and should not be construed as representing the policy or opinions of any agency of the United States government, or any institution in the Province of Quebec, public or private.

[1] Milton C. Weinstein and William B. Stason, *Hypertension: A Policy Perspective*. Cambridge: Harvard University Press, 1976. This study focuses on the United States. The introduction draws heavily on this source.

infarction); kidney failure because of arteriosclerosis of renal blood vessels is possible. All these conditions are accompanied by significant risk of death and considerable morbidity.[2]

The objective of treatment for essential hypertension is to lower arterial blood pressure. Currently the administration of antihypertensive drugs constitutes the main method of treatment, although adjunctive efforts to facilitate management and to reduce other risk factors for cardiovascular disease (such as high serum cholesterol and cigarette smoking) may be of additional benefit. Clearly much of this treatment involves physician-patient contact, which is the principal source of information about the medical process available for our analysis. The major component of the medical cost for treating hypertension is medication. Although the cost of medication is substantial, prescribing it is relatively routine and is generally not a separate reimbursable medical procedure in either the U.S. or in Quebec.[3]

The analysis presented below addresses the question of the relationship between the local physician supply and the choice of treatment mode. This question is central to much of the current policy discussion regarding the expected dramatic expansions in the U.S. physician supply and other issues such as national health insurance. The analysis is based on data from the Universal Health Insurance Plan in the Province of Quebec.

Hypertension was selected for this analysis because of the substantial range of physician discretion in determining an appropriate treatment regimen. It is apparently medically and professionally acceptable to see such patients anywhere from less than once to several times a year.[4]

Much of the variance is probably due to the nature and condition of the particular patient. Nonetheless, substantial anecdotal evidence suggests that physician discretion is great and that the wide range of treatment is not subject to professional review.

The analysis below provides empirical estimates of how much, if any, of the variance in the amount of physician services for treating hypertension is due to the local physician supply. In other words, does the available supply of physicians in an area bear any relationship to the revisit rate or cost per year of treating essential hypertension patients?

Other conditions, such as patient and physician characteristics, obviously also affect the treatment mode. Our analysis, therefore, controls for physician characteristics such as age, sex, specialty, as well as patient age, sex, family income, and general health status apart from the presence of hypertension (such as the presence of diabetes). Since there is no out-of-pocket cost to the patient of a medical service in Quebec, the decision to have a patient return more frequently for care is primarily determined by the physician.[5] Moreover, hypertension apparently does not manifest a syndrome that might encourage the patient to seek more medical care to relieve pain, as, for example, arthritis would. The patient accepts the diagnosis and decides whether or not to follow the physician's advice. Patient initiation of more care is less likely than in many other chronic conditions, where the symptoms are much more obvious.

The rest of this paper consists of three sections. The data are discussed in the next section. The subsequent section compares costs and revisit rates for treating hypertension. The final section provides a basic econometric analysis of the treatment of hypertension in Quebec, the principal analysis of this paper, and a summary of the results.

DATA[6]

The Province of Quebec has had a free-for-service universal health insurance plan since 1971. The plan reimburses nearly all of the Province's physicians for services provided to the 6.2 million residents. The data for this project have been constructed by collecting utilization information from the medical claims system for a stratified sample of 150,000 beneficiaries covering the period 1971 to 1975. The sample was stratified by location (65 areas), age and sex (5 groups), income (2 groups: chronic low income and non-low income) and by year of the plan (3 years). From the basic utilization sample, a subset of all persons who had received a medical service for a diagnosis of essential benign hypertension was selected for the analysis presented below.

COST AND REVISIT RATES FOR TREATING HYPERTENSION: UNIVARIATE MEASURES

This section provides an introduction to the data by presenting cost and treatment modes for hypertension using estimated means and rates of change of these measures over time. We focus primarily on three

---

[2] *Ibid.*

[3] The cost of medication constitutes most of the medical cost of treating hypertension. Weinstein and Stason assume total treatment cost for hypertension of $157 to $411 per patient year, after the first year of treatment, depending on medication used. Medications are 50 to 80 percent of the medical cost; their assumed three physician visits per year are priced at $54, or 13 to 30 percent of the medical cost of treatment (*ibid.*, p. 82). Since 1972, persons on public assistance in Quebec have received free drugs. The rest of the population bears the cost of drugs directly, although some supplementary insurance for drugs and medications is sometimes provided by employers.

[4] Weinstein and Stason assume three visits per year after the first year, although no source or justification for this number is given. We were unable to find any recommended revisit rates in the literature.

[5] We say primarily because even with universal insurance, which pays all physician charges, the patient still bears the travel and time costs of receiving medical care.

[6] See "A Study of the Responses of Canadian Physicians to the Introduction of Universal Medical Care Insurance: The First Five Years in Quebec," by Charles Berry, J. Alan Brewster, Philip J. Held, Barbara H. Kehrer, Larry M. Manheim, Uwe Reinhardt, final report for contracts HRA-230-75-0166 and 0167, Mathematica Policy Research, Princeton, NJ, June 1978.

Table 1. Percent of Physician Procedures for Hypertension in Quebec for Selected Age-Sex Groups 1974-75[a]

	Physician Procedures for Hypertension		
Age-Sex Group	*Percent of Total Procedures*	*Rank*	*Per Patient Year[b]*
Female (37-43)	2.5	5	3.69
Female (47-53)	5.5	3	3.53
Male (47-54)	2.9	2	3.67
Female (58-61)	9.0	1	3.65

[a]Simple means from a sample of 13,000 beneficiaries (roughly equal numbers in each age-sex group) stratified by market area and income. Only services for ICDA diagnostic code 401 were considered hypertension.

[b]These procedures do not include any procedures the hypertensive sample may have received that are not coded as essential hypertension. Per patient year is the mean of the sample.

measures of the performance of the medical delivery system: number of procedures,[7] number of office visits, and total costs. These measures are on a patient year basis, and are disaggregated by the beneficiary's age and sex (4 groups), location (urban, suburban, non-metropolitan), family income (non-low and chronic low), and sample year.[8] The following questions are answered:

- What are the basic distribution and composition of the medical procedures delivered for the treatment of hypertension?
- As measures of equity in the medical system, what are the differences between income and location groups as to the procedures received for the treatment of hypertension?[9]
- How have these treatment measures of hypertension changed over time?

The relative and absolute magnitudes of procedures delivered for hypertension in 1974-75 for our selected age-sex groups are shown in Table 1. For females (37-43), hypertension accounts for 2.5 percent of all procedures received and is the fifth most frequent diagnosis; for females (58-61) essential hypertension is the most common diagnosis and accounts for 9 percent of all services received. Interestingly, the average number of procedures delivered per hypertensive is almost uniform across age and sex groups — at 3.5 to 3.7 procedures per year.

Selected statistics on the treatment of hypertension for both the first and last sample years (1971-72 and 1974-75) are shown in Table 2 for the non-low family income beneficiaries. We observe that differences in total cost per patient year, number of procedures and office visits, and cost per office visit among age-sex groups are not large. Total physician costs were roughly $17 per year in 1971-72 and increased (at 10 percent per year) to $22 to $26 per patient year in 1974-75.[10]

[7]Procedures and services are used interchangeably and refer to those medical services which are reimbursed according to a fixed-fee-schedule by the Régie de l'assurance-maladie du Québec (Régie). In this paper the Régie is sometimes referred to as the Quebec Health Insurance Board.

[8]For a description of how the sample was drawn, see Philip J. Held and Larry M. Manheim, "A Test of the Validity of Diagnostic Data From a Universal Health Insurance Plan: Hypertension in Quebec," Working Paper, Mathematica Policy Research, Princeton, NJ, August 1978.

[9]While measures of medical procedures, costs, and visits in the treatment of hypertension are one set of indicators of the performance of the medical delivery system, we should not be oblivious to the shortcomings of this approach, which is limited to the data collectd by the Quebec Health Insurance Board (Régie) in the payment of medical claims to the providers of medical care. The data which are the basis of this analysis are devoid of information on the frequency and type of medications prescribed, the quality and amenities of care, and the patient outcomes. But with due acknowledgment of these shortcomings, the data employed for this analysis are better in content, sample size, and sample design than most previous research had available.

[10]Throughout this analysis costs are restricted to those procedures covered by the Health Insurance Board (Régie) which exclude medication costs but do cover some costs for laboratory and x-ray procedures. They do not cover all these ancillary costs because, while these procedures are reimbursable by the Régie, they are also covered by another provincial agency in the hospital sector whose records were not available to this project.

Table 2. Selected Statistics for the Treatment of Hypertension in Quebec by Age-Sex Group, 1971-75 (Non-Low Income)[a]

		Age-Sex Group			
Variable	*Year*	*Female 37-43*	*Female 47-53*	*Male 47-54*	*Female 58-61*
Total cost per patient year	1971-72	$16.62	$16.34	$17.87	$17.00
	1974-75	24.51	22.22	26.00	21.88
	Change per year	13.7%	10.7%	13.2%	8.7%
Total no. of procedures per patient year	1971-72	3.07	3.04	3.04	3.20
	1974-75	3.71	3.57	3.68	3.48
	Change per year	6.4%	5.4%	6.5%	2.8%
Total no. of office visits per patient year	1971-72	2.09	2.41	2.22	2.53
	1974-75	2.47	2.49	2.21	2.45
	Change per year	5.7%	1.1%	-0.1%	-1.1%
Cost per office visit	1971-72	$5.48	$5.43	$5.28	$5.18
	1974-75	6.30	6.08	6.41	6.09
	Change per year	4.7%	3.8%	6.6%	5.5%

[a]Simple means. All cost data are **physician costs only**. Approximate sample size is 13,000 cases of hypertension. Change is the compound rate of change per year.

Table 3. Distribution of Office Visits for Hypertension for Females (58-61) by Year, Non-low Family Incomes, 1971-75

Office Visits per Patient Year	Distribution of Beneficiaries (%) 1971-72	1974-75
1	8.4	9.2
2	37.2	36.6
3	20.2[a]	19.5[a]
4	13.1	11.0
5	6.4	8.5
6	4.2	5.8
7	3.6	2.7
8	1.6	2.3
9	1.5	1.5
10	0.9	1.0
11+	2.8	1.9
Total	100.0	100.0
Sample size	1,373	1,422

[a]Median

Generally, the number of office visits per patient year was constant, but the cost per office visit was increasing at 4 to 7 percent per year.[11] Total costs per patient year increased at an annual rate of 9 to 14 percent, with females (58-61) having the lowest rate of increase (8.7 percent per year).

No precise comparable cost estimates for the U.S. are available. But the costs in Quebec of $25.00 per year (Table 2) are substantially below the estimate of $54.00 per year for physician visit costs assumed by Weinstein and Stason in their recent policy analysis of hypertension.[12] It thus appears likely that the cost per patient year, exclusive of medication, is substantially less in Quebec than in the U.S.

The distribution of office visits for hypertension is shown in Table 3. The range of office visits is quite large, although the vast majority (80 percent) of the beneficiaries had less than five office visits for hypertension. The median was roughly three in both years, and the distributions were relatively constant over time. We can thus conclude that there were no basic shifts in the number of office visits per hypertensive patient per year between 1971 and 1974. As a further measure of central tendency, Table 4 provides estimates of the mean and standard deviation of physician visits per year for Quebec and the U.S. While the age-sex groups are not precisely the same,[13] the Quebec estimates are very similar to the U.S. estimates. For 1974, the Quebec unweighted mean was 3.38 while the U.S. mean estimate was quite similar, at 2.9 to 4 visits per year. Note, too, the similar standard deviations in the U.S. and Quebec, roughly 2.6 in both locations.

Some of the basic statistics for females (58-61) are disaggregated in Table 5. We observe substantial differences. Particularly interesting is the higher number of office visits per year for low income persons, holding location constant. Also note the differences in rates of change between the two sample periods. Overall costs, number of procedures, and office visits were increasing fastest in the urban areas, particularly for the non-low income stratum. For example, while the

[11]An exception was for females (37-43), where the increase was 5.7 percent per year, compounded annually.

[12]Weinstein and Stason, 1976.

[13]Recall that Table 1 showed little variation in services per patient year across age-sex groups in Quebec.

Table 4. Physician Visits Per Patient Year for Hypertension for Selected Age Groups in Quebec and the U.S., 1974

Location of the Visit	Quebec Female (58-61) Mean[a]	Standard Deviation	United States Male & Female (45-64) Mean	Standard Deviation
Physician's office	2.61	2.59	n.a.	n.a.
Hospital	0.36	n.a.	n.a.	n.a.
Home	0.14	n.a.	n.a.	n.a.
Outpatient department[b]	0.23	n.a.	n.a.	n.a.
Consultation	0.04	n.a.	n.a.	n.a.
Total	3.38	n.a.	2.92 to 3.99[c]	2.40 to 2.70[d]

[a]Unweighted sample means of beneficiary utilization sample. By definition, persons had to have at least one *procedure* for diagnosis hypertension to be included. In all likelihood such persons also had an office visit which implies that person with hypertension but without a doctor *visit* in the past year would not have been included in the hypertension subsample.

[b]Includes emergency room and a physician's office located in a hospital.

[c]Lower number is the mean for persons who "ever had hypertension" (2.70) and "now have hypertension" (3.14), including those persons who had no doctor visits in the past year for hypertension. The higher number is the comparable mean for "ever had" and "now have" but excludes those with no visit. The exclusion of the no-visit category would make this estimate most comparable to the Quebec estimates. Source: "Hypertension: United States, 1974," *Advance Data*, November 8, 1976, U.S. Government Printing Office, Washington D.C., p. 9.

[d]*Ibid.*

Table 5. Number of Procedures, Office Visits, and Total Physician Costs Per Patient Year for Treatment of Hypertension for Females 58-61 Years[a]

Family Income Beneficiary Location	Number of Procedures per Patient Year 1971-72	1974-75	Change per Year[d] (%)	Number of Office Visits per Patient Year 1971-72	1974-75	Change per Year[d] (%)	Total Physician Costs per Patient Year 1971-72	1974-75	Change per Year[d] (%)
Non-Low									
Urban	2.69	3.58	9.9	1.92	2.31	6.3	$16.65	$26.40	16.4
Suburban	3.49	3.57	0.8	2.71	2.45	-3.3	19.27	23.00	6.0
Non-metropolitan	3.23	3.44	2.1	2.61	2.48	-1.7	16.59	20.70	7.6
Chronic Low									
Urban	3.37	3.51	1.4	2.42	2.42	0.0	$18.30	$24.44	10.0
Suburban	3.34	3.31	-0.3	2.46	2.56	1.3	17.99	21.55	6.1
Non-metropolitan	3.92	4.01	0.8	3.14	3.00	-1.5	19.01	23.11	6.7
Weighted by Income[b]									
Urban	2.76	3.57	8.9	1.97	2.32	5.5	$16.82	$26.20	15.7
Suburban	3.48	3.54	0.6	2.69	2.46	-2.9	19.14	22.86	6.0
Non-metropolitan	3.30	3.50	2.0	2.66	2.53	-1.6	16.83	20.94	7.5
Weighted by Income and Location[c]	3.13	3.54	4.1	2.40	2.43	0.4	$17.36	$23.43	10.4

[a]All data are for diagnosis hypertension (ICDA code 401). Total sample size is approximately 4,400 cases of hypertension.

[b]Assumes weights of 0.90 and 0.10 for regular and chronic low income, respectively.

[c]Assumes weights of 0.39, 0.23 and 0.38 for urban, suburban and non-metropolitan, respectively.

[d]Annual compound rate of change of the annual means.

number of visits per year (per hypertensive patient) was decreasing or relatively constant in surburban and non-metropolitan areas, it was increasing in the urban areas at 6.3 percent. In addition, beneficiaries with non-low income and living in urban areas started from a relatively low base (e.g., 1.9 office visits per person in 1971), meaning that the large increases were to some extent a "catching up." Total costs per year for persons of non-low income were the same in urban and non-metropolitan areas ($16.62) in 1971-72; but by 1974-75 the cost for beneficiaries located in urban areas had increased to $26.40 while the increase for persons in non-metropolitan areas was half as great, reaching only $20.70.

The bottom portion of the table provides means weighted by income and location for females 58-61

Table 6. Percent of All Office Visits for Hypertension Categorized as Complete and Major Complete by Family Income, Year, and When Physician First Became Active, Quebec, 1971-75[a]

Year	Physician First Active[b]	Family Income[c] Non-Low (%)	Chronic Low (%)	Weighted Family Income[d]
1971-72	1971 or before	9.8 (223)	9.2 (140)	9.7
	1972 or after	5.8 (13)	0.0 (5)	5.2
	All	9.6 (236)	8.9 (145)	9.5
1973-74	1971 or before	17.8 (228)	13.5 (161)	17.4
	1972 or after	33.7 (38)	31.8 (38)	33.5
	All	20.1 (266)	17.0 (199)	19.8
1974-75	1971 or before	25.2 (233)	20.4 (137)	24.7
	1972 or after	39.8 (61)	32.0 (39)	39.0
	All	28.2 (294)	22.9 (176)	27.7

[a]Based on a sample of 1,316 beneficiaries approximately evenly distributed in the four age-sex groups of female (37-43); female (47-53); male (47-54); female (58-61). Sample sizes are shown in parentheses and refer to the number of beneficiaries. Estimates for non-low and chronic low family income are sample means.

[b]Physician who provided the modal office visit for hypertension for the year. Active status was assumed to be not a medical residen[t] and received $20,000 or more for the year from the Regie.

[c]Non-low family income includes persons whose family was not receiving public assistance in 1977. Chronic low income includes persons who were receiving public assistance in 1977 and first received public assistance in 1972.

[d]Assumes weights of 0.90 and 0.10 for non-low and chronic low income, respectively.

years of age, thus giving us "estimates"[14] for the province as a whole. We observe that costs per patient year were initially greatest in the suburban areas ($19.14) but, by the end of the period, the cost per patient year was greatest in the urban areas ($26.20). The rate of cost increase was greatest in the urban areas (15.7 percent per year) and least in the suburban areas (6.0 percent per year).

The increase in the cost per office visit over the period represents a shift to a longer office visit. The shift from the least complicated examination (ordinary) to more complex examinations (complete and major complete) is shown in Table 6. In 1971-72 only 9.5 percent of the examinations were of the more complex variety. But by 1974-75 the proportion of exams that were complete or major complete (CMC) had increased to 27.7 percent of all examinations. This trend to more intensive visits was true of older physicians who were practicing in 1972, but is much more prevalent among the newer physicians who have started practicing since 1972.

Summary

The statistics of this section provide some insight into the process of medical care delivery for hypertensives, how it has changed over time and how it differs by patient location and family income. This insight is limited by the nature of the information available through the Health Insurance Board's payment of claims, which is restricted to claims data on visits, certain costs, and certain procedures.

Basic observations that can be made from the analysis of this section include:

a. Our three aggregate measures of utilization (number of procedures, total cost per year, and office visits per year) suggest that the treatment mode for hypertension does not differ substantially by income or location.
b. Total physician costs per patient year in Quebec are likely to be substantially less than comparable figures in the U.S., even though the number of visits per year in both places is reasonably comparable.
c. Costs, number of procedures, and physician visits are reasonably constant across the four age-sex groups that are the focus of this analysis.
d. While office visits per year have remained relatively constant, total cost per patient year has been rising at 10 percent. Cost per office visit has been rising at approximately 5 percent per year.
e. The increasing cost per visit coincided with (in fact is identical to) the increase in the length of the visit. The proportion of ordinary examinations was decreasing and the proportion of complete and major complete examinations was increasing between the two time periods.
f. Chronic low income patients appear to receive more office visits per year than the other hypertensive patients.
g. Cost per patient year for treating hypertension increased most rapidly for beneficiaries located in the urban areas and least rapidly in the suburban areas. The difference in increase between the suburban and non-metropolitan areas was small.

ECONOMETRIC ANALYSIS

The fundamental goal of this analysis is to estimate the empirical relationship between the local physician supply and the treatment of hypertension in Quebec between 1971-75, the first 5 years of universal insurance. Do we observe that, where physicians are relatively plentiful compared to the population, patients with hypertension are treated differently from hypertensive patients in areas where physicians are relatively scarce? If physicians are in relatively short supply, are there fewer return visits or are the visits provided more costly? If physicians are more numerous, is the cost per year or the number of visits per year high relative to areas where physicians are scarce?

There are two reasons why we might observe a positive relationship between the local physician supply and our measures of utilization:

- *Non-price Rationing.* Since there are no money prices paid by the consumer, it is possible that excess demand exists, particularly in areas where there are very few physicians. In such instances, it is likely that physicians might ration their time among competing demands. Such behavior would include shorter and less costly visits as well as having the patient visit the physician less frequently.
- *Physician Inducement.* Much has been said about the control over demand for medical care that physicians have by virtue of their role as advisers or agents for the patient.[15] This theory states that a physician can induce or generate demand for services by advising the patient to have various procedures performed or by having the patient come back more often for checkups and the like. This theory would suggest that a positive relationship between the supply of physicians and our measures of utilization is the result of physicians, in areas of relatively greater supply, prescribing more medical care to compensate for the relatively lower demand per physician.

The empirical analysis presented below cannot distinguish between these two explanations. Our goal is to estimate empirically how large, if any, is the net result of these possible esplanations.

[14] See Held and Manheim, 1978, for a discussion of the weighting problems.

[15] Frank Sloan and Roger Feldman, "Monopolistic Elements in the Market for Physician Services," and Uwe Reinhardt's comments on the Sloan-Feldman paper, "Parkinson's Law and the Demand for Physician Services." Both papers were presented at the Federal Trade Commission's *Conference on Competition in the Health Care Sector*, Washington, D.C., June 1-2, 1977.

Our unit of analysis is the individual beneficiary with three cross-sections of observations, covering three sample years. We use two measures of utilization related to the treatment of hypertension in Quebec.

- The total cost per patient year for procedures paid for by the Quebec Health Insurance Board for treatment of essential hypertension.
- The revisit rate per patient year for hypertension as approximated by the number of visits and consultations (in all locations) per year for a diagnosis of essential hypertension.[16]

Our principal goal is to determine how important the local physician supply is in explaining observed differences in measures of utilization. Other factors can also be expected to have important effects on utlization rates, in particular, family income and location of the beneficiary. In order not to confound our interpretation of the importance of local physician supply, we must therefore "control" for these other factors. They can be divided into three general areas:

- Ambient market conditions such as degree of urbanization, average family income, and proximity to a large urban or medical center which are likely to affect utilization.
- Physician characteristics which are generally thought to affect the treatment mode[17] are likely to be especially relevant in hypertension, given the relatively recent nature of its clinical significance. Physician characteristics of interest include specialty, age, and sex.
- Patient characteristics might affect treatment pattern or utilization in several ways. First, a patient's general health status is likely to affect the treatment of hypertension in that less healthy patients and/or patients with complicating ailments are likely to be treated differently than patients who are otherwise healthy. Second, a patient's personal non-medical characteristics such as age, sex, and family income are likely to affect utilization. These latter characteristics have been shown to affect the demand for health care and may be proxy measures of health status as well.

Our two measures of annual utilization for essential hypertension will be regressed on the characteristics of the patient and the physician, as well as a series of market parameters. We now discuss the variables in more detail.

Market Conditions

Local physician supply will be measured by two variables, the active[18] general practice physician to population ratio and the comparable measure for specialists. The division of physicians into these two categories is common research practice and is used to reflect the degree of specialization of their practice. Since the population age-sex mix differs from market area to market area, implying different demands for medical care, two measures of market conditions will be used in addition to the physician-population ratio. These adjustments are the percent of the population in the market area that is less than 6 years of age and the percent that is greater than 64 years of age.[19]

Average family income (1970 census data by area), percent of the population living in rural areas, and proximity to large urban areas are used to approximate market factors that affect demand for medical care: travel time, convenience, and amenities for receiving care.

Physician Characteristics

Since patients can see more than one physician in a year, the physician who provided the model (most frequent) office visit was considered the primary physician. Approximately 19 percent of the sample saw more than one physician for hypertension during a given year. Characteristics of the physician who provided the modal office visit (modal physician) that are controlled for in the analysis are age, sex, and specialty. Since hypertension is a relatively recently identified condition, we hypothesize that older physicians will not treat hypertension with the intensity that younger physicians will, seeing patients less often and providing less intensive visits.

The sex of the physician can be expected to affect utilization, since females generally provide longer office visits and are likely to have more time available than are their male counterparts.[20] As a consequence, we expect female physicians to see their patients more often per year and provide more expensive office visits.

The specialty of the individual physician is reflected in the distinction between general practice and specialist. Further subdivision of the "specialist" group (e.g., between internists and surgeons) was avoided in

---

[16] An earlier version of this paper had used office visits only as the second measure of utilization, which prompted a reviewer to suggest that all visits, regardless of location (office, outpatient, physician's office in a hospital), was more appropriate. This paper reflects the reviewer's suggestion, although the basic results were not altered by the change. The means of Table 8 provide a basic comparison of the difference. The more general measure of visits is higher than the number of office visits by 0.4 visits (2.8 vs. 3.2). Overall the R^2 in the regressions increased roughly by 0.08.

[17] See, for example, Mark Pauly, *The Effect of Medical Staff Characteristics on Hospital Cost*, mimeo, Center for Health Services and Policy Research, Northwestern University, Evanston, Illinois, June 1977. Also see Beverly C. Payne and Thomas F. Lyons, "Method of Evaluating and Improving Personal Medical Care Quality, Episode of Illness Study," and a "Method of Evaluating and Improving Personal Medical Care Quality of Office Care Study," reports done for the Hawaii Medical Association by the University of Michigan, mimeo, February 1972.

[18] A physician was considered active if he/she was not a medical resident and received at least $20,000 for the year from the Quebec Health Insurance Board.

[19] See Berry *et al.*, 1978, Appendix B.

[20] Based on the 1977 Survey of Quebec general practitioners, female GP office visits were 29.5 minutes while male GP office visits were 30 percent shorter at 19.6 minutes. Furthermore, when asked of their desire for increased patient load, 18 percent of the males desired a greater patient load. More than twice as many females (39 percent) wanted a greater patient load. Only 8 percent of the female GPs wanted a decreased patient load, while 16 percent of the males wanted a decreased patient load. See Barry *et al.*, 1978, Ch. IV.

the interest of simplicity. Such a breakdown would be more important for the United States, where general practice physicians are not only older than specialists but also are a much smaller proportion of the total number of physicians. In Quebec general practitioners make up almost half the active physician stock.

It is not clear *a priori* what effect the physician specialty has on the utilization measures. On the one hand, specialists are trained to use more diagnostic tests and *may* be more informed and concerned about the health implications of hypertension. As a result, the utilization measures for hypertension may be higher if the physician is a specialist. Utilization by a specialist may also be higher because they may tend to see more serious cases (although this would be somewhat mitigated to the extent that our methodology is able to control effectively for patients' health status). A counter view is that the utilization per patient year would be lower if the modal physician were a specialist. The proponents of such a view argue that the specialists are more precise in their diagnosis because of the effective use of diagnostic tests and know more precisely what methodology to use for treatment.

Patient Characteristics

Age and Sex

Given that a person has been diagnosed as hypertensive, we would expect older persons to have higher utilization. Whether males use more or less care for hypertension than females is less clear, although males in general have lower utilization rates than females.

Family Income

Although the cost of physician and hospital services is paid by the provincial government in Quebec, family income is likely to affect utilization in several ways. Generally, persons with higher income are presumed to have greater demand for medical care than low income persons. Furthermore, higher income persons are more likely to be able to pay travel costs to receive medical care. Offsetting effects, however, are the generally higher time cost to receive medical care that goes along with higher wage rates and is presumably correlated with family income. Since drugs are provided at no cost to low-income persons in Quebec, and since medication costs are the majority of the costs for treating hypertension, low income persons may have a higher demand for physician visits for hypertension than the non-low income population because the cost of the resulting treatment will be lower.

Health Status

We presume that a person's general health status affects his demand for medical care. Similarly, if we presume that such decisions are made by the physician, prescribed treatment for a given condition is also liable to depend on a persons's general health status. One major determinant of health status is age and sex of the patient. An additional indicator of general well-being is a person's entire medical utilization record for the year.

There are at least two analytical reasons to include a person's overall utilization record. The first is the issue of complicating ailments. By this we mean those conditions which, other things equal, are likely to complicate the treatment of hypertension. For example, a patient with diabetes mellitus and hypertension may possibly require more treatment and physician monitoring for hypertension than a hypertensive patient without diabetes. Therefore, if we observed high utilization for hypertension for this patient but were unaware of the diabetes we might be confused in our interpretation of average utilization patterns. This would be the case if, for example, such cases were always referred to specialists. We would observe the higher utilization and ascribe it to the physician's specialty when it should have been ascribed to the health condition of the patient.

The second reason to consider a person's utilization record relates to the fact that the Health Insurance Board only records one diagnosis even if the physician indicates more than one on the claim record. If the visit or service were given a diagnosis other than hypertension (such as varicose veins of the lower extremities), we would be observing below average utilization for hypertension, and might ascribe such behavior to the wrong cause. If we were to control for this condition in a regression, the sign of the coefficient on varicose veins would be negative, indicating that the varicose vein utilization did not lead the less utilization for hypertension but was obscuring the visit for hypertension.

Our approach to identifying these two conditions — complicating medical conditions and substituting conditions — is twofold. First, we sorted the beneficiary records for the essential hypertension sample and determined what other diagnosis occurred relatively frequently. The resulting list of conditions was reviewed by physician consultants, who then recommended a list of diagnoses to be considered as control variables in the regression of hypertension utilization. The second approach to constructing such a list of diagnoses was to have our physician consultants review a sample of records for persons with hypertension. The results of both the computer listing of frequent other conditions for hypertensives and the physician review of a sample of records provided the list of diagnoses that would be expected to affect the observed utilization rates for hypertension. An example of the computer-sorted other diagnostic conditions for hypertensive patients is shown in Table 7.

This list of diagnoses and services believed important in understanding an individual's utilization for hypertension was divided into two groups. The first set included those items judged either to complicate the treatment of benign hypertension or to indicate the

Table 7. The 30 Most Frequent Other Diagnostic Conditions for 850 Females (58-61) Who Had 4-10 Services for Hypertension in One Year (1974-75)

ICDA Code	Diagnosis	Rank	n	Percent of All Services
000[a]	No Diagnosis Given	1	902	6.19
712,713, 714	Rheumatoid Arthritis, Osteo Arthritis, & Arthritis Unspecified	2	588	4.04
250	Diabetes Mellitus	3	456	3.13
410-414	Ischemic Heart Disease	4	375	2.57
300	Neuroses	5	364	2.50
627	Menopausal Symptoms	6	289	1.98
728	Vertebrogenic Pain Syndrome	7	262	1.80
344	Other Cerebral Paralysis	8	224	1.54
277	Obesity Not Specified As Endocrine Origin	9	221	1.52
470	Influenza, Unqualified	10	220	1.51
454	Varicose Veins of Lower Extremities	11	181	1.24
490	Bronchitis, Unqualified	12	165	1.13
731	Synovitis, Bursitis & Tenosynovitis	13	141	0.97
536	Disorders of Function of Stomach	14	117	0.80
370	Refractive Errors	15	106	0.73
247[a]	Control Visit Following Accute Illness	16	104	0.71
466	Acute Bronchitis & Bronchiolitis	17	95	0.65
692	Other Eczema & Dermatitis	18	91	0.62
996	Injury, Other & Unspecified	18	91	0.62
535	Gastritis & Duodenitis	19	81	0.56
174	Malignant Neoplasm of Breast	20	75	0.51
574	Cholelithiasis	20	75	0.51
783	Symptoms Referrable to Respiratory System	21	74	0.51
757	Congenital Anomalies of Skin, Hair, & Nails	22	71	0.49
595	Cystitis	23	67	0.46
400	Malignant Hypertension	24	66	0.45
009	Diarrheal Disease	25	65	0.45
285	Other & Unspecified Anemias	25	65	0.45
375	Glaucoma	25	65	0.45
785	Symptoms Referrable to Abdomen & Lower Gastrointestinal Tract	25	65	0.45

[a]Not an ICDA code.

severity of hypertension. They were, in effect, proxy indicators of health status. Included in this set were diabetes, heart condition, hospitalization for any diagnosis, malignant hypertension, and a consultation for hypertension.[21] The implications of the second group of conditions were unclear. An ad hoc method, therefore, was adopted to test whether they changed the basic results in any substantial way. Included in this latter set were missing diagnoses, neuroses, and treatment of varicose veins.

Empirical Results

The definitions of the variables used in the analysis and their means and standard deviations are presented in Table 8. The four sets of basic regression results are presented in Tables 10 through 13. Each contains the same two dependent variables, the cost per patient year, and the revisit rate per patient year, as approximated by the number of visits and consultations per year. The regression number is indicated at the top of each column. A directory of the independent variables used in each regression is contained in Table 9.

The regression results are presented in stepwise fashion, starting with basic patient characteristics, including age and sex, health status and family income, in Table 10. In the next set of regressions we test for the effect of the local physician supply, in Table 11, followed by the characteristics of the modal physician, in Table 12. Finally, in Table 13, we add sample year binaries, interaction terms to test for changes over time, and some measures of patient medical utilization not previously tested. The results will be discussed in order.

As a first impression we test for the effects of a patient's health status on the cost of treating hypertension. We observe large and statistically significant effects on the cost of having been hospitalized during the year. The effect of hospitalization (for any diagnosis) is $14 when the mean cost for hypertension is only $21. The effect of having at least one service for malignant hypertension is $18. While the effect of having had a service for heart disease is statistically significant at the 10 percent level, the

[21] A consultation is a relatively costly procedure compared to most other visits. While the ordinary examination cost $5, and the complete and major complete examinations cost $8 and $20, respectively, a consultation cost $15 to $30, between 1971 and 1975.

Table 8. Variable Definitions, Acronyms, Means and Standard Deviations for Regression Analysis of the Treatment of Hypertension, Quebec, 1971-75

Acronym	*Definition*	*Mean*	*Standard Deviation*
Dependent Variables			
TOTAL$H	Total cost per patient year for all procedures for essential hypertension ($)	20.619	22.946
OFFCVNH	No. of office visits for hypertension	2.750	2.797
VISITS	No. of visits and consultations for hypertension in all locations	3.162	3.670
Independent Variables			
F37-43	1 if beneficiary is female (37-43); 0 otherwise	0.112	0.315
F47-53	1 if beneficiary is female (47-53); 0 otherwise	0.305	0.461
M47-54	1 if beneficiary is male (47-54); 0 otherwise	0.108	0.311
LOWINCOME	1 if beneficiary is chronic low income; 0 otherwise	0.385	0.487
HOSPITAL	1 if # hospital[a] visits 0; 0 otherwise	0.123	0.329
HEART	1 if # procedures for heart disease[b] > 0; 0 otherwise	0.091	0.287
DIABETES	1 if # procedures for diabetes[c] > 0; 0 otherwise	0.062	0.241
MALIGNANT	1 if # procedures for malignant[d] hypertension > 0; 0 otherwise	0.026	0.160
PERCENT LT6	% of market area population less than 6 years old	10.174	1.217
PERCENT GT65	% of market area population greater than 65 years old	6.343	1.740
SPECPOP	(active specialists) / (pop x .001) in market area	0.298	0.321
GPPOP	(active general practitioners) / (pop x .001) in market area	0.378	0.089
MODLSPEC	1 if modal physician is a specialist; 0 otherwise	0.122	0.327
MODLAGE	age of modal physician (years)	43.364	11.998
MODLAGESQ	square of age of modal physician	2024.369	1116.121
MODLFEMALE	1 if modal physician is female; 0 otherwise	0.015	0.120
RURPOP	% of MA[e] population living in rural area	35.623	24.029
DISTMSQ	1 if market area is within a 50 mile drive of Montreal, Sherbrooke or Quebec City; 0 otherwise	0.677	0.468
INCMEAN	Mean 1970 household income in market area	7835.431	2171.072
YEAR2	1 if sample period is July 1973 — June 1974; 0 otherwise	0.352	0.478
YEAR3	1 if sample period is July 1974 — June 1975; 0 otherwise	0.347	0.476
LOWSPEC	Interaction of chronic low income (LOWINCOME) and modal physician specialty (MOLDSPEC)	0.036	0.186
INCYR2	Mean 1970 household income x year 2 binary	2767.379	3971.232
INCYR3	Mean 1970 household income x year 3 binary	2709.341	3927.246
LOWYR2	Low Income binary x YEAR 2 binary	0.140	0.347
LOWYR3	Low Income binary x YEAR 3 binary	0.139	0.346
CONSULT	Number of consultations for essential hypertension	0.042	0.200
MISSDIAG	Total cost of procedures with missing diagnosis ($)	13.017	39.120
INFLUENZA	Total cost of procedures for influenza unqualified ($)[f]	1.015	4.010
NEUROSES	Total cost of the procedures for neuroses ($)[g]	4.139	23.052
VVEIN	Total cost of the procedures for varicose veins of lower extremities ($)[h]	2.445	22.484
HOSPMISS	Hospital visits binary (HOSPITAL) x cost of procedures for missing diagnosis (MISSDIAG)	5.012	30.380
NUMBER OF OBSERVATIONS		13,177	

[a]For any diagnosis.
[b]ICDA codes 410, 411, 412, 413.
[c]ICDA code 250.
[d]ICDA code 400.
[e]MA is market area.
[f]ICDA code 470.
[g]ICDA code 300.
[h]ICDA code 454.

Table 9. Directory of Regression Results

Dependent Variables	*Regression Number*	*Table Number*	*Independent Variables*
1. Cost per year Visits and consultations per year	(1) (2)	10	Basic patient characteristics including health status and family income
2. Cost per year Visits and consultations per year	(3) (4)	11	Add to the above: measures of the local physician supply
3. Cost per year Visits and consultations per year	(5) (6)	12	Add to the above: characteristics of the modal physician
4. Cost per year (full sample) Visits and consultations per year	(7) (8)	13	Add to the above: market area characteristics, sample year, time interactions with area income and with low income status; certain other medical utilization indicators.

coefficient remains small ($1.20). Low family income did not affect costs in any appreciable way, nor did the age and sex of the patient.

Regression (2) shows the effects on visits per year. We find that, contrary to the results for *costs,* the age and sex of the patient affect the *revisit rate,* although the magnitudes are small. Consistent with the cost picture presented above, malignant hypertension has a substantial effect on the revisit rate (2.8 visits per year with a mean of 3.2). Given the nature of malignant hypertension,[22] such an observation is not a surprise.

In Table 11, we introduce measures of the market area physician supply, primarily the general practice to population and specialist to population ratios. The effect of a high general practice to population ratio on both costs and visits is positive. For costs, the coefficient of the ratio of GPs to population (GPPOP) has very small standard errors, meaning very precise measurements (t statistic = 6). For visits, the effect of GPPOP is statistically significant at the 0.01 level. It is also numerically large with an elasticity of 0.23; that is, for a 10 percent increase in the GP per population ratio, costs per patient year for treatment of hypertension would increase 2.3 percent.[23] The comparable cost elasticity for the specialist to population ratio is 0.06; that is, a 10 percent increase in the specialists per population ratio would increase costs by less than 1 percent. While the effect of both GPs and SPs is to raise costs as their numbers increase, the effect on the number of visits is different. The number of visits per year *decreases* as the number of specialists per population increases, but the number of visits *increases* as the number of GPs increases.

[22] Malignant hypertension is a more rare form of hypertension than essential hypertension, wherein the blood pressure is "extremely high."

[23] This is evaluated at the mean.

Table 10. Regression Estimates of the Cost and Visits Per Patient Year for Treatment of Hypertension, Quebec, 1971-75

Regression No.	(1)		(2)	
	Dependent Variables			
	Cost per Patient Year		*Visits per Patient Year*	
Independent Variables	*Coefficient*	*(t statistic)*	*Coefficient*	*(t statistic)*
F37-43	-0.230	(0.356)	-0.161	(1.547)
F47-53	-0.341	(0.758)	-0.077	(1.065)
M47-54	0.510	(0.780)	-0.182^{a}	(1.729)
LOWINCOME	-0.342E-2	(0.000)	0.265^{b}	(4.124)
HOSPITAL	14.301^{b}	(23.710)	2.100^{b}	(21.670)
HEART	1.199^{a}	(1.742)	-0.089	(0.808)
DIABETES	-0.752E-1	(0.095)	-0.233^{a}	(1.794)
MALIGNANT	18.093^{b}	(14.841)	2.830^{b}	(14.550)
constant	18.353		2.810	
Standard error	22.210		3.568	
R square (adjusted)	0.063		0.055	
Joint F statistic	111.943^{b}		96.720^{b}	
Mean of dependent variable	20.619		3.162	
Number of observations	13,177		13,177	

Statistical significance indicated as follows: a, 0.10 level; b, 0.01 level.

Table 11. Regression Estimates of the Cost and Visits Per Patient Year for Treatment of Hypertension, Quebec, 1971-75

Regression No.	(3)		(4)	
	Dependent Variables			
Independent Variables	Cost per Patient Year		Visits per Patient Year	
	Coefficient	(t statistic)	Coefficient	(t statistic)
F37-43	-0.147	(0.228)	-0.161	(1.549)
F47-53	-0.222	(0.495)	-0.085	(1.172)
M47-54	0.526	(0.806)	-0.170	(1.622)
LOWINCOME	-0.140	(0.032)	0.272[c]	(4.227)
HOSPITAL	14.354[c]	(23.816)	2.084[c]	(21.489)
HEART	1.087	(1.582)	-0.072	(0.653)
DIABETES	-0.228	(0.283)	-0.231[a]	(1.777)
MALIGNANT	18.149[c]	(15.015)	2.820[c]	(14.498)
PERCENT LT6	0.447[a]	(1.862)	0.042	(1.099)
PERCENT GT65	0.147	(0.963)	0.023	(0.954)
SPECPOP	4.004[c]	(5.699)	-0.265[b]	(2.345)
GPPOP	12.561[c]	(5.588)	0.994[c]	(2.748)
constant	6.900		1.932	
Standard error	22.157		3.566	
R square (adjusted)	0.068		0.057	
Joint F statistic	80.532[c]		66.177[c]	
Mean of dependent variable	20.619		3.162	
Number of observations	13,177		13,177	

Statistical significance indicated as follows: a, 0.10 level; b, 0.05 level; c, 0.01 level.

Table 12. Regression Estimates of the Cost and Visits Per Patient Year for Treatment of Hypertension, Quebec, 1971-75

Regression No.	(5)		(6)	
	Dependent Variables			
Independent Variables	Cost per Patient Year		Visits per Patient Year	
	Coefficient	(t statistic)	Coefficient	(t statistic)
F37-43	-0.274	(0.428)	-0.183[a]	(1.760)
F47-53	-0.263	(0.591)	-0.095	(1.321)
M47-54	0.396	(0.612)	-0.178[a]	(1.697)
LOWINCOME	0.235	(0.592)	0.254[c]	(3.947)
HOSPITAL	13.769[c]	(23.017)	2.071[c]	(21.320)
HEART	0.570	(0.837)	-0.076	(0.683)
DIABETES	-0.245	(0.307)	-0.238[a]	(1.833)
MALIGNANT	18.532[c]	(15.478)	2.833[c]	(14.568)
PERCENT LT6	0.264	(1.107)	0.036	(0.927)
PERCENT GT65	0.146	(0.962)	0.026	(1.049)
SPECPOP	2.556[c]	(3.620)	-0.237[b]	(2.066)
GPPOP	13.354[c]	(5.972)	0.867[b]	(2.387)
MODLSPEC	8.380[c]	(13.853)	-0.059	(0.605)
MODLAGE	-0.151[c]	(9.230)	-0.012[c]	(4.344)
MODLFEMALE	6.229[c]	(3.858)	0.074	(0.281)
constant	20.605		2.543	
Standard error	21.942		3.563	
R square (adjusted)	0.086		0.057	
Joint F statistic	83.222[c]		54.412[c]	
Mean of dependent variable	20.619		3.162	
Number of observations	13,177		13,177	

Statistical significance indicated as follows: a, 0.10 level; b, 0.05 level; c, 0.01 level.

Table 13. Regression Estimates of the Cost and Visits Per Patient Year for Treatment of Hypertension, Quebec, 1971-75

Regression No.	(7)		(8)	
	Dependent Variables			
Independent Variables	Cost per Patient Year		Visits per Patient Year	
	Coefficient	(t statistic)	Coefficient	(t statistic)
F37-43	-1.095[a]	(1.842)	-0.254[b]	(2.484)
F47-53	-0.642	(1.554)	-0.136[a]	(1.919)
M47-54	-0.925	(1.538)	-0.303[c]	(2.929)
LOWINCOME	1.978[c]	(2.845)	0.540[c]	(4.513)
HOSPITAL	8.103[c]	(12.895)	1.561[c]	(14.444)
HEART	0.128	(0.202)	-0.112	(1.033)
DIABETES	-0.697	(0.940)	-0.260[b]	(2.037)
MALIGNANT	17.088[c]	(15.372)	2.705[c]	(14.149)
PERCENT LT6	0.249	(1.076)	0.070[a]	(1.756)
PERCENT GT65	0.321[a]	(1.715)	0.033	(1.037)
SPECPOP	-2.154[b]	(2.115)	-0.324[a]	(1.849)
GPPOP	6.744[c]	(2.878)	0.640	(1.587)
MODLSPEC	2.634[c]	(3.843)	-0.578[c]	(4.904)
MODLAGE	-0.408[c]	(4.006)	-0.036[b]	(2.067)
MODLAGESQ	0.322E-2[c]	(2.949)	0.305E-3	(1.622)
MODLFEMALE	3.907[c]	(2.592)	-0.024	(0.095)
RURPOP	-0.512E-2	(0.319)	-0.126E-2	(0.458)
DISTMSQ	0.658	(1.386)	0.215[c]	(2.627)
INCMEAN	0.481E-3[b]	(2.411)	-0.689E-5	(0.200)
YEAR3	3.571[b]	(2.108)	0.524[a]	(1.798)
YEAR2	-0.357	(0.212)	-0.227	(0.780)
LOWSPEC	-2.483[b]	(2.103)	-0.023	(0.110)
INCYR3	0.134E-3	(0.658)	-0.360E-4	(1.031)
INCYR2	0.318E-3	(1.560)	0.293E-4	(0.835)
LOWYR2	-1.080	(1.182)	-0.288[a]	(1.832)
LOWYR3	-2.077[b]	(2.272)	-0.416[c]	(2.644)
CONSULT	41.745[c]	(44.127)	3.419[c]	(21.016)
MISSDIAG	—	—	—	—
INFLUENZA	0.136[c]	(3.057)	0.023[c]	(2.965)
NEUROSES	-0.019[b]	(2.453)	-0.417E-2[c]	(3.100)
VVEIN	-0.013	(1.604)	-0.231E-2[a]	(1.698)
HOSPMISS	0.028[c]	(4.288)	0.380E-2[c]	(3.374)
constant	19.296		2.580	
Standard error	20.340		3.498	
R square (adjusted)	0.214		0.092	
Joint F statistic	116.868[c]		43.890[c]	
Mean of dependent variable	20.619		3.162	
Number of observations	13,177		13,177	

Statistical significance indicated as follows: a, 0.10 level; b, 0.05 level; c, 0.01 level.

In the next set of regressions, (5) and (6) in Table 12, we introduce the individual characteristics of the physician who provided the modal office visit. We observe three effects on costs per year (all statistically significant): the cost per year *increases* if the modal physician is a specialist, younger than average, and female. Sex and specialty of the modal physician has no or little effect on the revisit rate. The revisit rate is lower for older physicians. The effect on the coefficients of the aggregate physician supply (GP and SP to population ratios) of introducing physician characteristics into the regression is to lower the positive cost effect of the specialist to population ratio, but to change the cost impact of GP to population ratio very little. Therefore, the effect on costs of the GP to population ratio remains relatively high (E = 0.24) while the elasticity of costs with respect to the specialists to population ratio falls from a relatively low value of 0.06 to 0.04.

Finally, in regressions (7) and (8) of Table 13, we introduce additional market area characteristics and medical information on the beneficiary. The overall effect of these additions is to lower the impact of some of the previous independent variables. For example, the effect of the GP to population ratio on the cost per year is reduced by half (from \$13 to \$6.8; elasticity = 0.12). The cost effect of the specialist to population ratio is changed from \$2.55 to -\$2.02. The effect of modal physician being a female is reduced from \$6 to \$4.

The reduction by a half of the elasticity (ε) of cost with respect to the GP to population ratio (0.24 to 0.12) is substantial. While it is not clear from the data of Table 13, this change appears to be primarily the result of adding the two year binaries. In other words, when

controlling for year (YEAR 2 and 3) in regression (7), the estimated effect of the GP to the population ratio on costs is half what it is when the year binary is excluded. This suggests that over time physician billings are increasing, but that the effect of the relative physician supply per se is not large. Recall for example that the elasticity of cost with respect to the specialist to population ratio also decreases when the year binary is introduced, but that the numerical value was originally fairly small ($\varepsilon = 0.04$).

It might be argued that the year binary is merely a measure of the increasing physician to population ratio and that the inclusion of both year and the relative physician supply is redundant. This argument, however, ignores several reasons for including both measures simultaneously. First, over time other characteristics such as highway improvements and personal income are likely to affect demand for medical care. Secondly, the long-run elasticity of demand is likely to be greater than the short run. Certainly learning how the system operates (for both provider and beneficiary) is likely to increase over time. Finally, it should be observed that the simple correlation of YEAR 2 and 3 with the relative physician supply (not shown) is relatively low and never exceeds 0.30. Therefore, it has been possible to estimate the separate effects of the year and physician supply variables.

While the positive coefficient on the YEAR 3 binary ($3.57 in regression (7)) by itself suggests physician-induced demand as the physician to population ratio increases over time, the inclusion of the relative physician supply as a regressor in the equation is ignored. Therefore, the increased cost over time remains unexplained in this analysis, but the increased cost over time does not suggest that the relative physician supply affects the costs of treatment. At least the coefficients of Table 12 suggest that the effect for GPs is not very large and that for specialists, the effect is negative.

It may be that the time effects are in fact a result of increasing physician supply for the province as a whole. That is, the Quebec Province medical community may change its view of the optimal treatment mode in response to changes in the relative scarcity of physician time. However, if this is happening, it occurs through a complex dynamic mechanism since the time effects cannot be explained by local area physicians adjusting their output to local supply. Also, since much of the increase in costs is due to a change in the type of visit supplied, it is possible that this time effect is solely due to a learning how to use the fee schedules, independent of physician supply.

We have also added in regressions (7) and (8) a term for the square of the modal physician's age (MODLAGESQ). The negative sign for the coefficient on physician's age and positive sign on the age-squared term describe a quadratic of the cost per patient year decreasing with physician age, reaching a minimum at 63 (see Figure 1). While the curve is mathematicallly constrained to turn upward at 63, there are in effect few active physicians older than 63. Consequently, the curve falls through the relevant range of physician age. The effect of physician's age on visits per year, while statistically significant, is small.

Chronically low income (LOWINCOME) persons receive more medical care per year for hypertension, as indicated by the positive coefficients on both cost and visits. The magnitudes of the additional treatment for the chronic low income persons are not trivial, 10 percent (1.98/20.62) of the mean cost and 14 percent (0.5/3.5) of the mean visits. The interaction terms of low income with years two and three (LOWYR2 and LOWYR3) show that the low income differential is decreasing over time; that is, the higher cost per year for low income persons was approaching the mean for the population as a whole.

While hypertensives in high income areas (INCMEAN) appear to receive slightly more treatment per year as measured in costs, the difference, while statistically significant, is very small. The revisit rate in high income areas is about the same as in low income areas. The effect of market area income does not seem to change over time as indicated by the statistically insignificant and small coefficients of the interaction terms between years and area income (INCYR2 and INCYR3).

Two other area regressors are entered as independent variables in regressions (7) and (8). The percentage of the population in the market area living in rural areas (RURPOP) and a binary indicating proximity to Montreal, Sherbrooke and Quebec City (DISTMSQ) are measures of travel time to a physician as well as other indicators of the convenience of urban life. Both of these coefficients in the cost equation are statistically

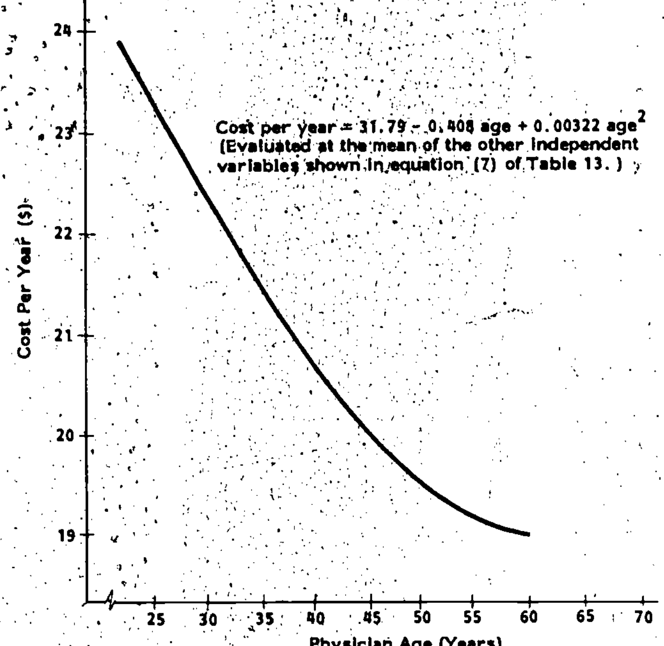


Figure 1.
The Effect of Physician Age on the Cost Per Patient Year of Treating Hypertension in Quebec, 1971-75

insignificant. Apparently, much of the effect of these measures is already reflected in other variables such as area family income (INCMEAN) and the specialists to population ratio (SPECPOP). There is a modest (6 percent of the mean) effect on the revisit rate of proximity to urban area, which presumably reflects some travel costs.

One issue of equity in medical utilization is whether low income persons see a specialist for hypertension at the same rate as the non-low income population. While we do not have a definitive answer to this question, the means of Table 8 suggest that they do. Since 39 percent of the sample is low income (LOWINCOME) and 12 percent of the sample had a specialist as modal physician (MODLSPEC), if low income persons saw a specialist at the same rate as the non-low income population, 5 percent of the low income persons (.39 x .12) would have a specialist as a modal physician. The means of LOWSPEC (an interaction term between LOWINCOME and MODLSPEC) suggest that 4 percent of the low income persons had a modal specialist, not too different than the hypothesized 5 percent. But the regression coefficient of the interaction term for modal specialist and low income (LOWSPEC) in Table 13 suggests that, even if low income persons see specialists at the same rate as non-low income persons, the services received per year are less by $2.48, not an insignificant amount compared to a mean of $20.

The very large coefficient ($41.75) on the consultation for hypertension binary (CONSULT in regression 7), with a very small standard error (t = 44.13), leaves little doubt of the large effect this measure has on average costs. We do not have a complete explanation of this phenomenon; it is probably a measure of case severity.

The three additional measures of patient's medical utilization (dollars per year for INFLUENZA; NEUROSES; VVEIN) are not easily interpreted. They are entered into the regression, however, in the context of controlling for multiple diagnoses (see above).

Finally, note that the interaction of the hospital admission binary and the cost of procedures with missing diagnoses (HOSPMISS) suggests that the missing diagnoses are for procedures delivered in the hopsital.[24] Results, not shown, which entered only the costs of procedures with missing diagnoses, showed a large coefficient in the cost equation. The interaction effect, as presented in regression (7), suggests that the cost of missing diagnoses was really an effect of hospitalization. Inclusion of HOSPMISS in regressions (7) and (8) controls more adequately for a patient being hospitalized and can, therefore, be interpreted as a general health status term. The consequence is that the procedure will provide better estimates for our variables of primary interest, since to exclude this term would in effect be omitting an important measure of health status. (This presumes, of course, that a patient who was hosptialized during the year is, on average, in poorer health.)

Conclusion

Holding patient age, sex, family income, location, and some proxy measures of health status constant, we can report the following regarding treatment of hypertension per patient year.

The effect of the local GP to population ratio is to increase cost a small to moderate amount (elasticity of 0.12). The effect of the local supply of general practitioners on the revisit rate is also positive with a rather low elasticity of 0.08, although it is not statistically significant under the usual criteria (t = 1.59).

The effect of the local specialists to population ratio is to reduce cost and the revisit rate, although the magnitude of the effect is very small (elasticity of -0.03).

Holding constant the local supply of physicians, costs are slightly higher, but the number of visits per year is lower if the physician is a specialist.

Both the cost and revisit rate decrease with physician age.

The costs are higher for female physicians compared to male physicians although the revisit rate is about the same.

Low income persons receive more services per year and have a higher revisit rate.

The differences between low income persons and the rest of the population are, however, decreasing over time.

Costs are rising over time, although the revisit *rate* is relatively constant.

Throughout this discussion we have maintained a neutral posture with respect to the implications of changing costs for the treatment of hypertension. Hypertension is undoubtedly a major health problem which does respond to treatment. When we observe that costs are increasing, it implies that services to patients are increasing. Whether the cost to society of these extra services is worth the benefits is a question clearly beyond the scope of this effort.

The results presented here show that treatment for hypertension is not independent of the local physician supply. The cost of treatment increases as the number of general practice physicians increases, but decreases as the number of specialists increases. The magnitude of both these effects, however, tends to be fairly small. Interestingly, the change in costs is associated more with the cost per visit than with the number of visits.

As was pointed out in the beginning of this section, a positive relationship between physician supply and treatment cost could stem from two different sources: non-price rationing and induced demand. But the results presented here suggest that whatever the source, the effect is not large, which implies that the issue is not as serious as one might at first suppose. These results are, of course, specific to one diagnostic condition and cannot readily be extrapolated to all medical care. It is

[24] See Held-Manheim, 1978, regarding missing diagnosis.

noteworthy, however, in that hypertension is a medical condition which allows considerable discretion to the physician in determining treatment mode and patient initiation of care is considerably less than in many other chronic conditions. It would therefore appear to be a natural candidate for testing the hypothesis of physician inducement. The results presented here suggest that if physician inducement is present, the magnitude is not large.

ACKNOWLEDGMENTS

The results presented here were originally part of a larger project that provided a more extensive analysis of the first five years of universal health insurance in Quebec. The report of the original project is titled: "A Study of the Responses of Canadian Physicians to the Introduction of Universal Medical Care Insurance: The First Five Years in Quebec," and was written by Charles Berry, J. Alan Brewster, Philip J. Held, Barbara H. Kehrer, Larry M. Manheim, and Uwe Reinhardt. The entire report is expected to be available from the National Technical Information Service (Washington, D.C.) in August 1978.

The cooperation of the Fédération de Médecins Omni-praticiens du Québec and the Régie de l'assurance-maladie du Québec are gratefully acknowledged. The opinions, however, are solely those of the authors.

The analysis presented here benefited greatly from the advice and consultation of Claire Bombardier, M.D., and Edward F. X. Hughes, M.D. The research assistance of Dale Schwartz is gratefully acknowledged. The typing was done in a most professional manner by Anna Mortimer.

A Market Model of the Distribution of Dentists

Gerald L. Musgrave
Hoover Institution
Stanford University

ABSTRACT □ An economic model is presented where the concentration of dentists is an integral part of the supply and demand for dental services. The three equation simultaneous system of demand, supply, and concentration shows how the quantity of dental services delivered, fees paid, and distribution of dentists are jointly determined.

Previous research concentrated on either non-economic criteria for practice location or single estimations, where the practice location decision was not part of an economic system. These studies used data on a national or state basis rather than on local or regional economic markets.

The results of this study seem to indicate that local and regional markets are an appropriate area for economic analysis. Our tentative results are that market forces do in fact determine the supply, demand, and market concentration of dentists. We find that the United States' experience concerning the distribution of dentists is not unlike the distribution of dentists and other health care providers in many other nations with diverse economic systems.

It appears that policy alternatives to alter the distribution of dentists must consider the impact of those policies on the entire dental market. Conventional policy recommendations have conventional economic results. For example, increasing the number of trained dentists will result in dentists locating in areas of economic need.

INTRODUCTION

The distribution of dentists in the United States is currently an important subject of health policy. In fact, the distribution of all health care providers has become a major policy issue not only in America but in many other nations. A caricature of the conventional wisdom seems to be that dentists choose their practice locations with little or no concern for the economic demand for their services. Then once the location decision has been made, the dentist can "adjust" his fee or the amount of service he provides to yield the income level he deems appropriate. In effect, what is being argued is that either conventional economics does not apply to dentistry or that conventional economic analysis must be substantially modified in the case of health professions.

One modification to conventional analysis was the introduction of the target hypothesis. Here it was assumed that the health providers did not attempt to maximize their own economic well-being but had a target level of economic affluence. They choose practice locations near their friends or families or near where they went to school. If the community provided many customers, they could pick and choose between patients, work short work hours, and establish fee schedules below what they might otherwise charge. On the other hand, if the community did not provide enough patients to yield the target income, then prices would have to be increased for the existing patients. Another tactic could be to induce the existing patients to consume enough services to generate the desired

income. The only constraint on the level of fees or quantity of services provided would be the target or aspiration level (possibly tempered by peer pressure) of the provider. (Here we have to be careful to refer to all health care providers and not single out dentists, since most research focuses on physician behavior.)

The last example has come to be called the supply induced demand hypothesis. If the hypothesis were true, we would find that location decisions were made independently of economic criteria, and those areas with relatively more dentists would have higher fees than areas with relatively fewer dentists. Apparently conventional economics does not work here either. We have a new health economics where more competition leads to higher prices rather than lower prices. One reason why health providers are allegedly able to escape conventional market forces is due to the public's medical ignorance or their blind trust in the physician or dentist. In these cases supply and demand influences the demand for medical services. Sellers of the service acting as agents for their patients are in effect making the decision to purchase the service. The net result is that demand for health services is higher where there are more providers. This is in contrast to conventional economics where the demand for a factor of production, here the input of labor service, is derived from the market demand for the product.

The major subject of this paper is to examine the extent to which dentists are in fact insulated from market forces. We question whether the United States' experience with its "maldistribution" of dentists is a unique result attributed to the lack of market controls. Also we attempt to determine what part market forces play in the location or per capita distribution of dentists. Finally, we show that the structure of the dental industry is in fact a result of market forces.

If the distribution of dentists is a result of market forces rather than being independent of those forces, any public policy designed to alter the distribution of dental manpower resources must be concerned with very broad policy impacts. A public policy that directly increases the number of dentists in low income or rural areas, would necessarily impact on the whole dental market. For example, one way to increase supply would be to reduce restrictions to or to promote entrance into the dental profession. The production of relatively more dentists would have the necessary consequences of lowering dentists' incomes from what they otherwise would have been. Removal of market restriction and regulations that discourage competition would also promote a more uniform distribution of dentists. However, if dentists are not subject to the regulation of market forces, then increased governmental intervention and regulation would be in order.

In the sections that follow we will review some of the existing literature that is related to these issues. Then we explain the economic model and data we use to investigate these issues. The final section is our empirical results and a summary of our findings.

LITERATURE REVIEW

The literature of interest is related to two issues: the supply and demand for dental services and the location of dentists. Since research in the area of dental economics is still relatively young, we have a very limited number of studies about the economics of dentistry. It is therefore appropriate to briefly consider other economic research on physicians and medicine in addition to that concerning dentists and dentistry.

The supply induced demand hypothesis can apparently be attributed primarily to two studies of physician service. Fuchs and Kramer [1] found support for the physician created demand hypothesis using data from both a time-series of national data 1948-1968 and a cross-section of state data for 1966. In the time series results they found that increases in medical technology and the number of physicians were primary factors in explaining the consumption of and expenditures on physicians' services. This result is primarily based on the positive relation between the quantity of medical services consumed and the number of physicians. Here the authors justified their position by noting the increase in consumption in the 1956-1966 period compared to the 1946-1956 period. They believe that the increase in the number of physicians over the period was an exogenous event. Also, they state that the increase in medical consumption was not due to either a movement along the demand schedule resulting from an increase in supply, or to shifts in the demand for physicians services due to changes in income, insurance or demographics.[1] These hypotheses are not tested using statistical methods.

Since utilization was thought to be influenced by technological change, Fuchs and Kramer examined a 1966 cross-section of data. Here the effects of technical change would be held constant by the methodological design. It was assumed that the average vintage of human and non-human capital were identical across the 33 states in the sample. The analysis developed in this study is much more sophisticated in that a simultaneous four equation (and two identities) model was developed. The most important equations for our purposes are the ones for per capita general practitioner visit equivalents and the physician/population ratio. These equations were defined as:

(1) Per capita visits — f[average price or net (net of insurance benefits) price, per capita insurance benefits, Physician/Population ratio, median income, per capita hospital beds].

(2) Private physicians per 100,000 population — f(average price, output per physician, number of medical schools in the state, per capita hospital beds, median income).

The third equation relates output per physician to average price, number of physicians, and hospital beds per capita. The fourth equation is an insurance

[1] Fuchs and Kramer [1], pp. 16-17.

equation. Here per capita insurance benefits are determined by per capita consumption of physician visits, average price, union members per 100,000 population, and the ratio of health insurance premiums to benefits.

In the first equation is the primary source of the physician induced demand information. Eleven equations in total were estimated on the basis of the variables cited above. In any of these equations, at most three variables were included in any one run. The statistically significant results are, in general, that average or net price is negatively related to the number of patient visits per capita. The remaining variables (median income, per capita insurance benefits, hospital beds per capita and physicians per 100,000 population) are all positively related to the number of patient visits per capita. The significance of the physician/population ratio was attributed to the physicians' ability to increase consumption rather than a supply induced movement along a demand curve. This interpretation was based on the fact that physicians' fees were in the equation. Fuchs and Kramer attribute the positive sign of the coefficient to the fact that as physician density increases, the average travel time falls as does average waiting time. Thus the real cost falls if fees are held constant. However, the authors state that because of the low price elasticities of demand (none larger than -0.36), this influence appears to be too small to account for the size of the density elasticity (0.507-0.335). Recall that the coefficient is based on physicians per 100,000 population. The authors also reject the notion that an excess demand exists for physicians in general or that intra-state mobility restrictions are due to licensing regulations. They provide no explanation, but they note that Feldstein [2] proposed this explanation for his time series results. Fuchs and Kramer attribute Feldstein's results to omitting technical change in his study. Since Fuchs and Kramer are assuming that technology is the same across all states, the influence of technical change has been held constant. The final explanation is that physicians create demand. Fuchs and Kramer conclude that physician induced demand is more important than the influence of income, price, or insurance coverage. They also conclude that the physician density ratio is positively related to fees, medical schools, per capita hospital beds, educational, cultural and recreational facilities. The quantity of service per physician is negative related to the number of physicians and does not rise with higher fees. The final equation concerning medical insurance, unlike the quantity of physician service, does not appear to be sensitive to differences in median income. In addition, per capita insurance benefits are statistically significant when related to unionization and the price of insurance.

Many economists might find this model difficult to understand. The equations are not clearly related to conventional supply and demand models. The authors do not discuss the economic forces in the context of individual, firm behavior, or market behavior. The results are purely based on empirical findings. This, of course, does not mean they are incorrect. However, it remains to be shown that the Fuchs and Kramer results are or are not consistent with conventional economic theory.

The Feldstein [2] paper cited by Fuchs and Kramer [1] should be discussed for a number of reasons. First, this study used time series data from the 1940-1966 period. The period was chosen for data availability and because 1967 was the first full year that medicare and medicaid programs were in operation. Also, this model is interesting because it considers the simultaneous influence of various economic factors on price and physician supply. Feldstein also introduces a dynamic price adjustment relation. Feldstein differentiated between the net price that consumers pay after insurance compensation and the average price the supplier would receive. He formulates a per capita demand equation and a supply equation where the supply response is a function of the ratio of physician services per physician. When the demand equation was estimated the net price elasticity was positive. The explanation for this was ". . . that the observed price quantity combinations do not lie on (or around) the demand function but that, at observed prices, there is excess demand." The concept of excess demand is consistent with the notion that a shortage of physicians existed in this period, (see Arrow and Capron [3], Fein [5], Rayack [4]). Feldstein introduced a price adjustment mechanism and two reduced form equations (price and supply) were estimated. The physician/population ratio was considered as exogenous as it was not statistically significant in either equation. The price elasticity of supply was negative and the implied price elasticity of demand was still positive. Feldstein considers these contradictions of conventional economic theory to be consistent with the excess demand hypothesis. If the excess demand hypothesis were true, physicians have restrained price increases and/or constricted service output because of their substantial discretionary power to vary price and quantity. It might be noted again that Feldstein does include a time trend in the supply equation to account for technological change.

One major result of this study is a clear statement of the excess demand hypotheses and that the result is the ability of physicians to have discretionary power over the prices they charge and the quantity of services they supply. However, the results are not supportive of the supplier induced demand hypothesis.

Another study, this one by Newhouse [6], has been quoted as providing evidence for the supplier induced demand hypothesis (see Evans [7], p. 192). The objective of the Newhouse paper was to develop an analytic framework to distinguish between competitive and monopolistic market structures for physician services. Newhouse concluded that the market was monopolistic. However, Frech and Ginsburg [9] found an inconsistency in the Newhouse model that made the

statistical tests inconclusive. Newhouse had assumed a constant marginal cost for the monopoly model and increasing marginal costs (for increasing output) in the competitive case. When Ginsburg and Frech reformulated the models they found the relation between price and number of physicians to be identical in both the monopoly and the competitive models and thus the two models are really the same. In addition, they indicate that the mobility of physicians was not adequately modeled and that the location response of physicians may be a major factor in the labor market adjustment to changes in demand.

A later study by Evans, Parish, and Sully [8], using Canadian data, investigated the impact of group practice on output per physician and discussed the extent to which physicians were able to generate demand for their own services. They found that group practice had little impact on medical service output and that physicians appeared to have an impact on the demand for their services. However, Evans, et al. were not able to differentiate the demand generation hypothesis from an alternative one that postulates a backlog of unmet existing needs which were being supplied by the existing physician stock. Another possible explanation of these empirical results, in addition to the demand creation hypothesis, is related to the total price that a patient pays. These prices include the search costs, travel and waiting time. As the density of physicians increases, these implicit prices are reduced.

The results of the studies cited above are quite different from the previous literature. In the pre-1970 literature one generally finds that models were based on profit maximization and a major concern was an examination of the extent of monopoly power. Two studies in the mid-1950's representative of this approach are Kassel [10] and Hyde and Wolff [11]. The classic study on the monopoly gains accruing to members of the medical profession is by Friedman and Kuznets [12].

However, the few studies that purport to show supply induced demand are now accepted as the convention wisdom. Apparently, many policy makers believe that the market for medical services is fundamentally different from other markets, (see Dyckman [13]). And because of the acceptance of these results, conventional economic policies including the stimulation of competition are considered to be inadequate and possibly perverse. Policies to promote an increase in the supply of physicians or dentists would not lead to an increase in the amount of health services or a lowering of the price of health services. The necessary governmental policy would apparently be price controls, regulation of practice location, restriction of provider discretion, and other "direct" actions.

But there are numerous other studies that do not accept the view that the medical profession is insulated from market forces, that it can create its own demand, or that its members can earn whatever target income they choose. Returning to the Feldstein [2] paper, we now note a criticism of it first indicated by Brown and Lapan [14]. The criticism is that the positive price elasticity obtained by Feldstein in the estimated demand curve results from deficiencies in defining the price variable and that the supply estimates are biased because of identification problems. Based on estimates of their own, Brown and Lapan conclude ". . . our findings, which are exactly opposite to Feldstein's, lead to different policy statements, and are consistent with standard economic theory." Feldstein [15] rejected the criticism on a number of grounds. One interesting remark in the reply was Feldstein's argument that physicians have special latitude in many practice decisions and are apparently not fully constrained by market forces. He thus rejected the Brown-Lapan service queue or excess demand hypothesis. Here Feldstein says, "There are a number of ways in which doctors determine the type of cases and persons who receive care: physicians choose their specialty and location; specialists choose what types of cases to accept on referral; general practitioners decide how many house calls to make, which patients to treat, which to refer and which to discourage from seeking further care." He then buttressed this with "There is substantial sociological literature on this subject (Friedson [16], Martin [17], Reder [18])." Much of Feldstein's argument could be applied to other occupations and industries such as retail sales, barbers, cosmeticians, auto or T.V. repair. One wonders how this argument differentiates medical services from other services. The point here is not an attempt to arbitrate the debate but only to indicate that these matters are far from settled.

A more recent paper by Steinwald and Sloan [19] investigated a number of influences on physician fees. The authors did not formulate a model but they attempted to determine what economic market conditions would be consistent with their findings. Their data were from mail questionnaires returned in the fall of 1971 by members of the American Medical Association. They considered a number of explanatory variables and their equations might be thought of as alternative reduced form equations. The most important result for our purpose is the estimated relation between fees and the physician population ratio. Steinwald and Sloan found that demand variables exerted definite and expected impacts on fees, and they found that mark-up pricing theory was an inferior explanation of physician pricing behavior when compared to profit maximization. The negative coefficients of the physician population ratio variables in the fee equation were cited as evidence against the target income (via supply induced demand) hypothesis. However, it should be noted that these results were not uniform. For general practitioners (using country data) and for general surgeons (using state data), the results were consistent with traditional economic theory. Internists, Ob-Gyn, and pediatricians had opposite results. Since the first two categories are the most numerous, the results in these cases were judged to dominate the other results. If non-price rationing were

important in medical decision making, then the empirical results would not have resulted in as many findings that were consistent with conventional economic theory as actually occurred. Steinwald and Sloan conclude that their results are sufficiently consistent with the standard profit maximization model that future work should be based on this assumption rather than on supply induced demand, target income, or mark-up pricing hypotheses. (Although it should be noted that mark-up pricing can be consistent with profit maximization.)

Turning to dental markets, the most recent study is by Kushman and Scheffler [20] where they formulated a monopoly based on Newhouse [6], Newhouse and Sloan [21], and Frech and Ginsburg [9]. The major focus is on deriving and estimating a price function for dental services. They estimate a single reduced form equation for each of five separate procedures and a combined price index. The results of the five service prices (Prophylaxis, Two-Surface Amalgam, Single Extraction, Porcelain Bonded Jacket, Complete Upper Acrylic-Base Denture) are based on questionnaire responses from individual practitioners. In all of the equations, hygienists' earnings, per capita income, percent of population are positively related to fees. Fees are negatively related to the percent of population on fluoridated drinking water. Kushman and Scheffler found lower fees were charged by older dentists. The density of dentists in the state where the individual dentist practices was included in the fee regression and it was statistically significant as were the other factors and its sign was positive. The authors attribute the result to the possibility that travel time could be lower where dentists are relatively more plentiful. This is also consistent with the Holtmann and Olsen [22] result that increased waiting time (another implicit price like travel time) exerted a negative influence on the demand for dental visits. In that study travel time was only statistically significant in one of their four equations. This result is consistent with findings of other researchers who have investigated what might be called amenities or qualities of physician visits. Acton [25] found that non-monetary factors became important determinants of demand when out-of-pocket money prices were low. As governmental health legislation or private health insurance expands, these issues will be more important. To measure these effects Acton analyzed "free" outpatient departments in New York City. It was found that travel time functioned as a price in the demand for "free" services. The computed elasticities of distance were equivalent to price elasticities computed by other researchers. Also, Acton found that there was substitutability between the private and public sector physician services based on these non-monetary travel time and distance factors. Sloan and Lorant [26] found positive correlation between length of visit and physician density and that waiting time was negatively correlated with the physician density ratio [27].

Returning to Kushman and Scheffler's final macro equation, we note that because of potential data problems and measurement errors, the five equation model was collapsed to a single equation where a fee index was the dependent variable. A number of results were presented attempting to measure the effect of dentist age on fees and the results were again generally consistent with the disaggregated equations. The exceptions are that hygienists' earnings are not statistically significant in the fee equation and the impact of age on fees is substantially reduced. Both income and dentists per 1,000 population are positively related to fees and fluoridation is negatively related to fees. A final section of the paper dealt with the issue of comparing the results of this estimation to other studies. Kushman and Scheffler find their parameter estimates to be consistent with the estimates of Shepard [23] and Maurizi [24], if one accepts the monopoly assumption. In defense of the monopoly assumption, a competitive market model was formulated that could be compared to the assumed monopoly model. Under these conditions one requirement is equality between the coefficients of the price variable in the linear supply equation and the linear demand equation. Kushman and Scheffler do not estimate supply and demand and cannot perform the test directly. However, they argue that since their results are close to those of Shepard and Maurizi, coefficients from those studies can be used. Using those results the competitive market alternative is rejected.

The Maurizi [24] study mentioned above provides an interesting if somewhat flawed view of dental markets. Chapter three is an attempt to estimate the costs of a national dental insurance program. In that attempt a model of the demand and supply of dental services was constructed, using the 1962 American Dental Association *Survey of Dental Practice* and state data. The fee and treatment data on individual practitioners is essentially the same as that reported in Kushman and Scheffler. The demand equation relates the quantity of dental services to price of dental services, consumer income, and the extent of fluoridation. Unfortunately, we do not know how the price of dental service was constructed. Presumably it is an index of five fees that were sampled in the ADA questionnaire. The source and measurement of the remaining variables is also unclear.

The supply equation relates quantity of dental services to price of dental service, number of chairs in the office, number of full time auxiliaries, and the number of hours worked by the dentist. The exogenous variables in this equation seem more representative of those found in production functions rather than supply functions. A specification issue may be present here, since the relations between the aggregate production function, the reduced form cost function, and the resultant supply function are not explained. The method used to estimate the parameters was said to be two stage

least squares. However, the author only indicates one reduced form equation and states that the model "... is exactly identified: that is, there are as many unknown variables (two) to be solved for the model as there are equations" (page 61). Actually both equations are over identified. Two stage least squares would be an appropriate estimation technique and assuming the other problems to be minimal we can look at the results. All of the coefficients are statistically significant. The elasticity for price in the demand equation is negative (-1.76) and positive (0.79) in the supply equation in conformance with traditional economic theory. The income elasticity of demand is positive (1.06) and the coefficient for fluoridation is negative. As expected, the number of chairs, full time auxiliaries, and hours of work by the dentist all have positive signs.

Another part of the study was an estimation of the impact of regulations on the hourly earnings of dentists. A single regression equation was estimated relating hourly earnings to years of labor force experience, the square of labor force experience, years of education, number of active dentists per 100,000 civilian population in the state, per capita income in the state, average failure rate on state dental license exam in the 1960-1969 decade, and a binary variable for self-employment status. A sample of 420 dentists was drawn from the 1970 Census Public Use Sample (1 in 100 sample) of the civilian labor force. The empirical results are that experience is positively related to dentists' income, as is education. Less restrictive regulations which allow dentists to delegate more tasks increases dentists, hourly earnings. Dentists in states with higher average failure ratios had higher earnings than those in states with an average failure ratio. The author indicated that self-employed dentists earned more than employee dentists. However, since the coefficient for that variable is negative in the regression equation, that interpretation is either questionable or a misprint of the sign occurred. The computed sign of per capita income was positive and the sign of the physician population ratio is negative. Neither of these variables is statistically different from zero, so we cannot take these results as support for the traditional view that increased competition, as measured by the dentist population ratio, leads to reduced hourly earnings.

Another paper that considered the demand for dental services is by Upton and Silverman [57]. They were interested in estimating the influence of income and fluoridation on the consumption of various dental procedures. Their data were obtained from 15 midwestern dentists for one week in 1966. Half of the towns had fluoridated water supplies the others did not. The following regression was run:

$$\ln(T_i) = \alpha + \beta \ln(I_i) + \gamma F_i + \xi_i$$

where ln(T) represents the natural logarithm of the number of treatments, I_i is the median income of the i city in 1960, and F is a dummy variable equalling 1 if the city has fluoridated water and zero otherwise.

Regressions of this form were run for 13 different procedures and the quantity of dentists as dependent variables. The "income elasticities" ranged from 0.37 for "extractions — children" to 6.11 for "inlays" and the income elasticity for dentists being 2.26. Extractions and inlays are in quotes because in actual practice situations such procedures can differ widely. We have found that care needs to be taken when defining dental procedures and in the data section we discuss the 1972 ADA coding system. Also we note that the equations that were estimated were actually Engel curves rather than demand curves. But because price data are difficult to obtain especially on a local basis many researchers regress utilization on explanitory variables when consumption patterns are under investigation. The Upton and Silverman results are plausible. They find that those procedures that are more costly in terms of time or materials or are subject to alternative treatment modalities such as crowns, inlays, and bridges have higher income elasticities than extractions and restorations. Restorations of diciduous teeth (children) had one of the highest income elasticities as one might expect. For all procedures the income elasticities averaged 2.39 which is higher than the 1.10 estimate of Friedman and Kuznets [12], 1.20 by Benham, Maurizi and Reder [33], 0.12-0.41 by Holtman and Olsen [22], 1.03 to 1.82 by Feldstein [57], and 1.06 by Maurizi [24]. However, the data used in this study are from a small market area and such results are not inconceivable in the absence of the other problems we mentioned.

Fluoridation was found to be a statistically important factor in reducing the demand for dental services. Their analysis indicated that the use of fluoride would reduce the demand for dental services by over 55 percent. This seems to be a rather high estimate since fluoridation's main impact seems to be on caries in childen. We also note that substantial adult tooth loss results from periodontal disease and this was not considered in the model although it is mentioned by the authors. It remains to be seen if other studies on a national basis confirm the Upton and Silverman estimates.

The most comprehensive study of the dental services industry was done by Feldstein [58] in 1973. That study was not limited to the supply and demand for dental services but also considered the impact of governmental financing on the provision of dental services, developed an econometric forecasting model for dentists, and analyzed the economics of dental education. Our brief remarks focus on Feldstein's work on the supply and demand for dental services. However, Feldstein's work more clearly than any other shows the complex, interconnected and simultaneous nature of the many divergent influences in the economics of dentistry.

Feldstein was careful to include both economic and non-economic determinants of dental demand. His analysis was particularly noteworthy in that he clearly recognized the economic basis of many factors that others thought of as completely noneconomic. The treatment of public attitudes toward dental care is a

good example of his insights. Many other researchers discussed the importance of attitudes and knowledge regarding dental care in the consumption patterns for dental services and measured these attitudinal differences by years of formal education received by the head of the household. Feldstein clearly explained the interrelations between education and income, race, urbanization and government subsidy programs. He went on to show the differences in dental consumption patterns as a function of differences in income. More extractions and dentures were consumed by lower income families and more preventive and maintenance work was demanded by higher income patients. Data were presented that allowed per capita visits to be crosstabed with income and education. Per capita utilization increased with education holding income constant, however the major difference in utilization was between income levels holding education constant. The analysis proceded to show how the changes in consumption patterns over time were related to income levels.

The analysis of price, income, and fluoridation as determinants of demand is very interesting. Using dental visits as the dependent variable, the own-price elasticity was estimated to be 1.43. The income elasticities were reported to vary depending upon the sample and the choice of dependent variable. Using national time-series data (1929-1970), the income expenditure elasticity is 1.7, indicating that a one percent increase in per capita income would yield a 1.7 percent increase in dental expenditures. A 1961 cross section of 38 cities yielded an expenditure elasticity of 1.82, and across states the estimate is 1.03. When visits rather than expenditures are used the estimated elasticity is 1.55.

Fluoridation was found to be an important factor in reducing the demand for dental services, especially in children. The results indicate that fluoridation has the effect of reducing dental expenditures by approximately $5.00 per person per year. This estimate is substantially higher than the $1.00 per capita estimate of Upton and Silverman [57].

Feldstein attempted to estimate the economic impact of public expenditures on dental health programs and dental services demand. No statistical relation was found. It was believed that the small amount spent, only a few cents per person, made the statistical estimation difficult. However, these payments are in fact concentrated on a specific income group and if data on this group were available more work could be done on this question.

Feldstein's analysis of supply is also one of the most lucid. He clearly differentiates between long-run conditions where increases in the quantity of dental services supplied results from increases in all factors or production and the short-run case where increases are primarily from longer work hours of existing personnel. The analysis primarily focuses on the long-run issues of increased supply due to changes in dental technology, greater utilization of auxiliaries, and improving organizational efficiency through group practice. It is clearly recognized that the supply of dentists is closely related to the demand for dental services. Also some time is spent in detailing the alternative ways dental services could be expanded including reducing regulations that inhibit mobility, developing alternative methods of training new dentists, employing more auxiliaries, and establishing various governmental subsidies.

The results of this section are that the long-run supply of dentists is in fact responsive to economic pressures, most especially the financial rewards of being a dentist. Also that if equality in dental services is a public policy objective, then recommendations should focus on increasing the consumption of dental services. Heretofore policy seems to have been aimed at increasing or redistributing dentists which is only one and, as Feldstein documents, an expensive way to increase dental services.

LOCATION

A logical place to begin a brief review of the physician/dentist location literature is the Friedman-Kuznets [12] study. The major objective of that study was to examine the determinants of income earned from independent professional practices. In addition to the professions of medicine and dentistry, studies were also conducted on law, certified public accountancy, and consulting engineering.

Two characterizing features of the practice location received attention: the size of the practice community and its geographic region. Here the size of the community is based on six to eight population groups and the regions are the nine major census regions. Friedman and Kuznets found that variations in community size had a much larger effect on variations of earned professional income than did variations in geographic regions. Differences in professional incomes across geographic regions were smaller than differences across community sizes. Also the differences in professional income attributed to these large geographic regions were found to be attributed almost entirely to regional differences in urbanization.

For the five professions in general, mean professional income increased with city size. The ratio of dentist income in cities of 1.5 million or more to dentist income in cities of under 2,500 people equaled 1.88. In these large cities, dentists earned incomes 120 percent of the national mean income in the United States; whereas, dentists in the small towns earned 64 percent of the mean of all reported incomes. Medicine was an exception to the general trend of rising incomes with larger community size. Physician incomes were higher in middlesize (10,000 to 1,499,999) than in the largest cities (ones with 1.5 million or more people). The coefficient of variation of incomes was much less sensitive to changes

in community size than was income itself. The results were not uniform, but the coefficient of variation seems to fall from larger to smaller cities. However, the smallest cities have more variation in professional incomes than do larger cities. No consistent regional differences in the variability of incomes was discovered.

In discussing the results of the study, Friedman and Kuznets argued that the absence of regional differences in average income indicated sufficient geographic factor mobility to prevent the establishment or maintenance of income differentials. Thus, differences in income across communities could not easily be attributed to factor immobility. This is because mobility among various size cities within a region must be at least as large as across geographic regions. The differences in income, it was argued, would have to be attributed to higher productivity of the professionals in larger cities or the non-pecuniary benefits of residing in smaller towns.

Friedman and Kuznets' results indicate that income distributions across community size were similar in character and magnitude for all of the five professions. In addition, the economic market forces that determined the geographic income distribution for physicians, dentists, lawyers, certified public accountants, and consulting engineers were similar to those determining the income distribution of the general working population.

It should be added that it was not argued that pecuniary economic forces were the only causes of the geographic distribution of professionals. They explicitly noted the importance of residing near family and friends. Having knowledge of an area, its climate, physical and cultural attributes, and availability of professional facilities were explicitly mentioned as influences in the location decision process. One point of including these factors in the location decision process was to explain why differences in money wages should occur. That is, we should not expect all dentists' incomes to be equal. Even if all professionals were equally productive (in a physical sense) and free to move without any barriers to relocation, we would not expect them to have equal incomes. Rather we would have "equivalent" incomes. Incomes would be "equivalent" in the sense that monetary differences would be compensating for generally observed differences in the non-monetary advantages or disadvantages of indifferent areas. However, we know that dentists in particular and professionals in general are not mobile. Pashigian [28] found dentists to be second only to judges as the most immobile professionals. These results are consistent with Wash [30] who found that turnover of health workers, in general, was significantly lower than that of the general labor force. As the skill level fell mobility increased. Dentists were the least liable to transfer occupations with physicians and pharmacists being more mobile across professions. Pharmacists are twice as likely to change occupations as are dentists, and registered nurses are almost three times more likely to change occupations. As Friedman and Kuznets note, the distributional changes generated from new entrants into the profession, rather than the mobility of persons already in active practice, are the most important influences in changing the regional distribution of professionals. The risk of beginning anew in a different location, plus the cost of new physical capital and the funds necessary to support the family in the interim period, are substantial disincentives to mobility. Moreover, those who are the least financially successful in their current location are least likely to have the financial resources to move. Successful practitioners probably have the capital but not the economic motivation to move.

Many other researchers have investigated the locational issues first raised by this path breaking study. Glass and Baldwin [29] administered a questionnaire to a sample of 1,244 dentists in New England who began practice between 1956 and 1965. They asked, "What single factor most affected your decision to choose the present location of your practice?" The following were the most frequently reported:

	Frequency (in percent)	
Location Factors	*Present*	*Future*
Shortage of Dentists	20	22
Good Living Conditions	15	14
Nearness to Family, Friends	8	1
Hometown	7	0
Availability of Office Space	9	19
Offer of Partnership	6	1
Quality of Practice	2	4
Economic Factor, High Fees	2	7

Clearly economic issues were important to the location decision. Nearness to family and friends and returning to hometown are all important to the initial decision. These factors were called "social" by the authors; however, they could also be interpreted as information sources and thus they are certainly economic. The possibility of the dentist being known in his home community is probably an important marketing influence. "Good living conditions" are probably close proxies to the non-pecuniary factors mentioned in the Friedman and Kuznets study. Glass and Baldwin also asked, "What single factor would be most important if you were to change your location in the future?" These responses are listed in column two in the previous table. Economic factors appear to grow in importance as practice experience grows. It is unexpected that office availability looms so large in the thinking of dentists. No mention of cost of office space was mentioned, so it is hard to tell what the economic logic of the response was. It may be that the potential for improving the practice income or service by improving the practice location only becomes evident after some years of experience.

The authors indicate that the availability of recreational facilities were listed "far more frequently" by dentists in Maine, New Hampshire, and Vermont. Availability of cultural activities were more frequently

listed by those in Connecticut. This reinforces the general idea that personal preferences are influencing the location decision of dentists simultaneously with direct economic issues.

Location decisions of recent dental graduates were examined by the American Dental Association [39]. They found that practices were located near parents', spouse's parents' residence, pre-dental college, or dental school. A small minority (5.8 and 2.4 percent respectively) reported their practice to be near the location of their advanced dental program or military assignment.

Recent graduates were asked to indicate the factors they believed to be responsible for their choice of practice location. The most popular response was "Geographic Preference" which sheds little light on the decision process. The next most frequent responses (in order) were; recreational and cultural advantages, type of people in locale, economic level of community, population to dentist ratio, influence of family and friends, climate, childhood residence, availability of dental facilities, influence of other dentists, and influence of spouse. These results are consistent with other research in that both direct economic forces and indirect or social factors are jointly determining the location choice.

Since changes in the local practitioner population ratio are likely to originate from new entrants into the profession, modeling this migration pattern could be important. Yett and Sloan [31] examined the migration patterns of new physician entrants into private practice in 1966. They found that the physicians' familiarity with the area in terms of "events" of birth, medical school, internship, and residency were important explanatory factors in determining initial practice location. In addition, other economic and demographic factors were also important. These factors were: climatic conditions, urbanization, failure rate on state licensing exam, mean net income of all specialists in private practice, per capita income of residents, and the change in per capita income in the period 1960-1966. Both familiarity with the general area, in this case the SMSA or state, and other demographic and economic considerations determine the distribution of physicians in this data set. Weiskotten et al. [32], using cross-tabulations on data from 1915 to 1950 found generally the same results concerning previous knowledge of and experience with geographic areas. Residency training was the most important factor in determining the state of practice. However, a trend away from large cities of 500,000 or more toward smaller cities of less than 25,000 was noted. Also they found a relatively higher percentage of graduates locating in states with higher than average per capita personal incomes. While two out of five medical graduates were practicing in a city the same size as the one that they resided in prior to entering medical school, that proportion steadily fell in the period. Specialists were concentrated in the larger cities, limited specialists were uniformly distributed in all but the small towns, and general practitioners were more densely distributed in the smaller communities.

These results are clearly what one would expect from an active economic market for medical services. This Weiskotten [32] study is the most often cited study in the medical literature of location; unfortunately, the references tend to focus on the schooling, internship and residency issues and exclude any mention of the economic or market forces that interact with the informational factors.

Another often cited study about migration and location of medical professionals is by Benham, Maurizi, and Reder [33]. This study incorporated four cross-sections of state data from 1930, 1940, 1950, and 1960. Because of data problems they were not able to pool the observations for time series analysis; nevertheless they performed several regressions using first differences. In general, they found the supply of dentists in a state to be related to the size of the population, growth in the population, total personal income, failure rate on licensing exam, physician or dentist income, and population in urban areas. Here again we find evidence that market forces of supply and demand for medical services are in fact determining the geographic distribution of health professionals. Benham et al. found that both dentists and physicians tended to locate in areas with higher effective demand for their services. They also found that this economic behavior was "superimposed" on location preferences that caused dentists and physicians to sacrifice money income for amenities in various locations. Dentist mobility was restricted more than that of physicians due to state licensing regulations.

Two studies by the same researcher shed some light on the market structure of dentistry. In the first study Tryon [34] examined the variations in the distribution of dentists among SMSA's using the 1960 census. This study and the previous ones did not attempt to formulate a formal behavioral or economic model and then develop testable hypotheses. The method was purely empirical. Various independent variables were chosen on the basis of their plausibility as measures of a metropolitan area's ability to attract dentists. Tryon concluded that the dental service industry depended on consumer purchasing power and motivation. Motivation in this context would be called preference by most economists and motivation was measured by educational level. As we mentioned in connection with Paul Feldstein's book, education and income are highly correlated and it is difficult to separate the income from the preference effect. Dentists were also found to be attracted to locations where purchasing power (per capita income) is greatest. This factor was predominant when the influence of population size was held constant. Here again we find dentists being driven by economic forces, like other service providers, to locate where the economic need is the greatest.

The second Tryon [35] study found different results. Here 49 census tracts within the Hartford Connecticut

Standard Metropolitan Statistical Area in 1970 were used. Tryon attempted to correlate the dentist population ratio within these census tracts with other available data. In general the zero-order correlation coefficients were not statistically significant. However, two variables were significant. Percent of economic activity devoted to business use and number of bus trips per day passing through the census tract both had significant (and positive) correlations with dentist density ratio. Other possible influences, such as socio-economic status of the tracts, residents, race, land use, and population density, were not related to the distribution of dentists. It is not clear how one would expect these influences to impact the distribution of dentists on the census tract level.

The most appealing interpretation of these results is that the sample space was inappropriate for the question under study. Rather than interpreting the results as being in conflict with other studies because of the urban nature of the data, it appears that we need a separate mini-micro theory of dental practice location across census tracts. However, the issue for conventional economic analysis is whether or not census tracts are economic markets for dental services. It appears that the results reject the notion that census tracts are in fact geographic markets for dental services. Also other studies, including the previous one done by Tryon, seem to support the idea that if large cities or SMSA's are examined, standard economic results are obtained. Also it appears that state or groups of states are useful when issues such as regulations or climate are under study.

Rimlinger and Steele [36] found per capita effective buying income to be the prime determinant of M.D.'s per 100,000 population. In this study, county groups were used as the relevant economic area. These groupings were stablished by the Public Health Service and published in the Health Manpower Source Book [37]. In general, one county is used if it contains a metropolitan area of more than one million, adjacent counties are grouped so that they have similar populations, semi-rural and isolated areas are combined to form generally contiguous county groups. The point here is that aggregate geographic areas larger than cities and smaller than states can be effectively employed to model the distribution of physicians. In this study, it was found that the physician population ratio was statistically related to fees and average income in the region. High income areas had substantially more physicians per capita than did low income areas. Also it was shown that the physician density ratio was an increasing (parabolic) function of per capita effective buying income. Another finding was that physicians did not reduce their work effort in areas of higher physician population density. The authors attribute this to lack of mobility between geographic areas. However, they found that desire for leisure is not a strong motivating force in physicians' choice of location. Finally they note that as geographical income differentials are reduced, the distribution of physicians would be more uniform. However, as we have seen, even if the distribution of money income were equal across all areas we would not have economically equivalent real incomes. This is because some areas are generally agreed to be more desirable than others and these non-monetary attributes will be reflected in compensating variances in monetary incomes in long run competitive equilibrium.

In a second study by the same two researchers, Steele and Rimlinger [37] using time-series data from 1950 to 1959 found that the degree of urbanization and increases in population were the two most important influences in determining variations in the percentage changes in the stock of physicians. The study was performed for all physicians, rather than only those in private practice. It is known, as indicated by Steel and Rimlinger, that substantial differences exist between the more inclusive and the less inclusive physician groups. In any event, the market forces operating in this market are clear and these authors make an important point. The point is that physicians seem to be *following* the general population trends.

In this time period it was found that the relative inequality of distribution of physicians based on per capita income of the areas residents decreased and at the same time the urban-rural inequality increased. However, it was the movement of the general population to relatively higher income areas that dominated rather than the movement of physicians themselves. Steele and Rimlinger argue that physicians followed the population trends but at a slower rate. Also they note that the population trend results in a smaller fraction of the population residing in geographical areas served by relatively fewer physicians. It appears that market forces are at work causing physicians to locate in concert with changes in general population trends. Those characteristics and conditions that cause relatively few members of the general public to live in rural areas (one characteristic of rural areas is the fewness of sellers of all goods and services) also cause few physicians to locate in these areas.

Ball and Wilson [38] also used a cross-section of county groups for 1962 and 1966 to explain the distribution of physicians and health facilities. Only the 1962 regression results were displayed and they were consistent with other studies using larger geographical areas. For all M.D.'s plus osteopaths, provider density was positively related to effective buying income per capita, the presence of a medical school, negatively related to population, and not related to percentage change in population. The population variables certainly have unexpected results. Growth areas appear to be less likely to immediately draw M.D.'s and osteopaths when compared to stable or declining county groups. The independent variables explained about 60 percent of the viariation in the M.D. density ratio. A second regression on G.P. densities was also run but with less success. The R^2 was 0.076. The sign for per capita income was positive; however, signs of the remaining variables were reversed from the first

equation. One difficulty may be due to the fact that G.P. density across these county groups shows little variation. The 1962 data were displayed for various size classifications of the county groups. Counties with a major population center had 34 G.P.'s per 100,000 residents and isolated rural county groups had 33 G.P.'s per 100,000 residents. The variation in density ratio was only evident when specialists (who concentrate in population centers) were included. Specialization in dentistry is not as significant as in medicine, so this influence is not expected to effect our work.

Virtually hundreds of studies have been done on the location of physicians, dentists, health care professionals, hospitals and other medical facilities. A detailed bibliography on physician location was published by the Robert Wood Johnson Foundation [40]; literature on dentists and physicians location was reviewed by Deane, McClung, and Tobin [41], medical geography was reviewed by Armstrong [42], Dever [43], May [44], and Shannon and Dever [45]; influences of race and urban conditions are reviewed in Elesh and Schollander [46] and Guzier and Jahiel [47]; and the economics of regulating the dental profession in Holen [51], MacBride [52], and Sheppard [53].

The final item of interest concerns the experience of other nations. Is the American situation, with our reliance on market forces, substantially different from other countries that rely on non-market administrative processes to allocate scarce medical resources? We have not made an exhaustive survey of this literature. However, it can be said that our distributional patterns are not unique. In nations as different as the Soviet Union and Great Britain, we find remarkably similar distribution patterns with similar difficulties in staffing rural health services see Duffand and Hollingshead [48] and Markovin [49]. Five separate studies of Australia, Belgium, Canada, Norway, and Poland were summarized by Roemer [50]. The basic economic system in each of these nations is different. The first four countries have various degrees of national health insurance. Poland has a National Health Service, funded from general revenues and all resource allocations under that system are directly administered by the government. Roemer explains that the common attribute of these health care systems is that the majority of the payments are financed collectively and the services are provided as public benefits. In all five countries, inequalities in the distribution of physicians and other health care providers were found between cities and rural areas. Roemer added parenthetically that this was the existing condition in "virtually all nations in the world." It was also found that rural residents did not seek treatment as often as urban dwellers, even if the direct fee was very low. These countries found that urban areas were preferred over others. As the number of health providers grew, the rural areas were easier to staff. Dentists in Norway and Poland increased relative to the total population and in those two nations the urban-rural distribution became more uniform. In most of these programs central planning or regional planning established general strategies for hospital construction and to some extent personnel programming. The result was regional centers for health facilities. From the discussion it appeared that these results were much the same as those that have resulted from market forces in the United States, with specialists in urban areas, general practitioners in rural areas and limited specialization in transitional regions in between. These five governments each promoted specialization in "general medicine" and in the United States a parallel development might be the "family practice" specialist.

In Poland, two years of mandatory rural service was required for medical school graduates from 1948 to 1963. However, with the growth in the number of physicians, the necessity for this policy vanished. Norway also had a similar policy of mandatory rural service which was abandoned with the expansion of physician supply. In reviewing the numerous actions of the regulatory authorities in providing more equal distribution of health providers, it became clear that government provided medical services are in fact subsidies to rural residence. It is as if the government were providing direct grants to the rural residents. The most remarkable result of these findings is that within systems that do not rely on market forces to allocate medical services directly we find results almost completely consistent with traditional market analysis. These nations found that increases in supply of medical personnel substantially affects the rural-urban distribution of providers contrary to what has become the conventional wisdom in America.

ECONOMIC FRAMEWORK FOR ANALYSIS

The economic model that is postulated is a three equation simultaneous system for equations. The first equation is the demand for dental services, the second equation is the supply of dental services, and the third equation is the location or the regional concentration of dentists equation. These three equations jointly determine the equilibrium quantity of dental services, the fees for those services, and the concentration or distribution of dentists in a market area.

We have chosen counties and country groups as the relevant economic markets for dental services. In our literature review we noted that because of data availability most previous studies focused on either time series over states or on the nation as a whole. Other studies used small geographic areas such as census tracts and some used SMSA's. SMSA's seem to be an appropriate geographic size for an economic model of dentistry. However, we could not investigate locational issues by restricting our observations to large cities or counties with a major population center. We need data that include a distribution of observations that include major poulation centers and smaller more rural areas. This is not simply a data issue but instead it is a central

ssue of our economic model. We maintain that the competitive market forces are operational on the local level. We would like to abstract from the other economic issues, such as technological change or interstate regulation, and focus on the competition in local markets. It should not be surprising that when time series data are used technical change is important, or that when cross-sections of states are analyzed that regulatory influences are important. We expect that if these influences are held constant by either statistical methodology or experimental design, that competitive market forces would be important. This is what we propose to do. We will model the dental market for one large diverse state, California, using cross-section data of counties and county groups. The data are specifically discussed in the following section. What is important for the analytics is our attempt to focus on regional competition by sample selection. We hope to extend the results presented here to include regulatory and other factors that vary from state to state.

Demand

The demand equation takes the form:

$$Q_{dj} = A_0 + A_1 P_j + \sum_{i=2}^{n} A_i X_{ij} + U_{dj} \qquad (1)$$

Where Q_{dj} is the quantity of procedures demanded in the j^{th} market. P_j is the mean fee paid in market j and the X_{ij} variables are the N-1 exogenous determinants of demand in the j^{th} regional market. U_{dj} is the stochastic error term in the market demand equation.

We expect the sign of A_1 to be negative indicating the customary inverse relation between price and quantity demanded when the other relevant economic variables are held constant. The exogenous variables in this demand equation were selected on the basis of having appeared in other research studies or as proxies for economic factors mentioned in the literature we reviewed. One would certainly prefer to develop the rational for selecting the variables from fundamental economic behavior assumptions. However, as mentioned in the introduction, this analysis is (as are the others we reviewed) in what Tukey would call the exploratory data analysis stage rather than tests of formal economic theory.

We [illegible] expect income to play an important role in determining the demand for dental services; in fact, expect the demand for dental services to be more highly related to income than other medical services. Medical services are generally covered by some form of insurance or government subsidy program, dental services are not. Dental care can often be postponed (even though this leads to expensive or disastrous consequences later) and medical services are less subject to postponement.

One price the patient must pay in addition to the fee itself is the transportation cost to the dentist. In areas where the population is more concentrated, the representative patient would be on average closer to the dentist and thus have a shorter implicit travel cost to pay. Along with the implicit prices that patients pay one would like to see the extent to which insurance coverage altered consumption. Some observers believe that recognizing the need for dental care and overcoming the expected fear of pain (both real and imagined) are in fact more important in dental consumption patterns than are price or income.

If this were true, we would find that price and income were not significant in the demand equation. Other factors such as education, IQ, and dental knowledge would be dominent forces. Another way to present the argument is that if economic factors do not play an important part in consumption patterns, then the consumption patterns of covered and non-covered individuals should be the same (assuming the two groups are matched on all other relevant criteria). Family budget studies would be the best way to approach the problem from this direction. Unfortunately, we do not have family budget data, nor do we have data on the number of people who have dental insurance by county. However, a large share of dental insurance is held by union members and we have included union membership as a proxy for insurance. We know that this measure will not satisfy many because of the importance of the issues. However, it was included on an experimental basis.

Finally, we include Medi-Cal payments for dental services. These payments are made to low income and new wealth individuals and their dependents on a need basis. Some payments are made to people declared "medically indigent" who would not otherwise qualify for government assistance. We expect that these payments would cause the demand equation to shift to the right. In effect dental demand would be augmented by the addition of Medi-Cal patients who would not receive treatment in absence of the subsidy program.

Supply

The supply function can be written as:

$$Q_{sj} = B_0 + B_1 P_j + B_2(\text{Dentist/Population}) + B_3 \text{Wages}_j + U_{sj} \qquad (2)$$

Q_{sj} is the quantity of dental procedures supplied in the j^{th} market and P_j is the mean fee charged. The dentist population ratio measures the competitive conditions in the local dental markets. We shall see in the next section that the dental density ratio, like quantity and price, are jointly determined. Wages received in the area are included as a cost of labor to the dental practice. We would like to have had the wages or incomes of auxiliary personnel or dental assistants. These data were not available nor were other operating expenses. This may not be a serious objection if in fact the cost of capital is not significantly different across the geographic areas.

Dental Concentration

The dental concentration function can be written in the following form:

$$(\text{Dentist/Population})_j = C_0 + C_1 \left(\frac{\text{Fees x Procedures}}{\text{Population}}\right)_j + \sum_{i=2}^{n} C_i X_{ij} + U_{cj} \qquad (3)$$

The dependent variable is the number of full time dental practitioners in the j^{th} county or region divided by the total population in that market area. The first variable on the right hand side is the mean fee charged multiplied by the number of procedures performed in the area divided by the population. It can be thought of as per capita total revenue or per capita total expenditures. As we indicated, only a small number of patients were covered by dental insurance and thus we do not have a differential between the patient's expenditure and the revenue received by the provider. We would expect dental concentration to be higher in those market areas where per capita dental expenditures were higher.

We would also expect dentists to prefer to locate in areas where personal incomes were higher. This is due to two effects; the first is similar to a derived effect. We would expect more dental services to be supplied in higher income areas and thus more dentists would locate in these areas due to natural market forces. This will be the derived demand effect in that the demand for dentists is derived from the public's demand for dental services. A second influence is the separate attraction of the area. We expect higher income areas to have more of the cultural and recreational activities that appear to be important to dentists.

Areas that are more economically prosperous have a number of institutional similarities. We attempt to measure the economic conditions in an area by the number of savings and loan institutions. These institutions are important financial intermediaries in California. They represent not only a source of funds for the active housing market in the state but they also represent an important savings bank function. In our model these institutions represent a measure of the economic health of the area. Areas with a large number of branches represent the fact that such areas are sources of funds and thus represent areas with higher discretionary income. As users of funds, these geographic areas can be considered areas of rapid economic development and are potentially attractive areas for new dental practices.

Union membership in the model is a proxy measure for the number of patients with dental insurance. We would expect the relation between dentist concentration and insurance coverage to be positive. In effect we have wage payments being paid to union members in the form of dental services. We would expect more dental services to be consumed under these circumstances than if the union members received the same dollar value of income in direct wage payments. Such economic conditions would result, other things held constant, in a stimulating effect on the location of practices. However, since this measure is an imperfect representation of insurance coverage, a number of other results are possible. It may be the case that union membership will measure socio-economic status or differences in the distribution of industries across the geographic regions. In these cases the expected sign would be indeterminate.

Population density might be an influence in the location decisions of dentists. We saw that the geographic density of patients could measure the time of travel to the dentist's office, and as such it influences the real price that is paid for dental services. It is also true that population density is a measure of urbanization of an area. Our review of the literature indicates that dentists would prefer less to more urbanization if this factor alone could be adjusted. However, locations usually can not be separated this way when the actual choices are made. In our model we can attempt to see what influence urbanization has on location. Our expectation is that fewer dentists per capita would be praciticing in locations with high population densities. The apparently counter intuitive result comes from our hypothesis that practice location is not exogenously determined but is the result of market forces generating the equilibrium between supply and demand for dental services themself.

The final issue of interest concern the migration of dentists. Apparently three competing hypotheses exist. First, population trends do not influence dentists location patterns. Second, dentists follow past population trends. Third, dentists anticipate population trends and establish practices in future growth areas.

Our savings and loan data could yield some insight if we interpret higher values to mean higher current growth rates. We also obtained data on past population growth and expected future growth. The past growth was for the census period 1960 to 1970 and the projections were from 1976 to 1985. Each growth rate was adjusted to account for differences in size of the market areas. The computed signs of the coefficients for these variables is expected to be positive.

The magnitude of the coefficient or its statistical significance should shed some light on the relative importance of these factors.

The model that is developed in this section attempts to show that dentist concentration is an integral part, in fact a result, of the supply of dental services. We have a three equation system of equations that determine the quantity of dental services performed, the fees that are charged, and the per capita distribution of practicing dentists. The market demand for dental services is a function of the fee for the service, patient ability to pay as measured income, and a number of other demographic influences. The supply of services is a function of the fee received by the dentists, the dentist concentration ratio, and the costs of providing the services. Dental concentration is determined by the interaction dental services supply and demand and a

number of outside influences including the economic and cultural conditions of the area and population growth patterns in the geographic area.

Data

Because this study focuses on county and regional dental markets in California, it is necessary to detail some of our data sources. Local and regional data are difficult to obtain and probably the most difficult data to obtain are dental fees and the number of procedures performed. The most useful source of price and utilization data is the Health Insurance Assocation of America (HIAA). These data are coded using the 1972 American Dental Association four-digit codes for dental procedures and the three-digit postal zip code for location. The data we used were the mean fee charge and the number of charges reported for the various procedures in the geographic market. This information is distributed semi-annually to members of the HIAA under the title *Dental Prevailing Healthcare Charges System*. These data represent the actual fees (normal, customary, and research) charged by those dentists who perform work for patients whose insurance company is a member of the Association. Dentists submit these fees to the insurance companies for reimbursement and the fees are inclusive of any co-payment. Thus, these fees are intended to be the prevailing fees as would have been charged if the patient were uninsured. Insurance carriers do audit work and there is some evidence that these fees are in fact representative of what would have been charged had the patient been uninsured. However, we do not know if any systematic relation exists between those dentists whose patients have insurance and others. Also we do not have any information concerning the treatment mix of uninsured versus insured patients. We have assumed that these data are representative of the dental market in each area. Our sample observations are for the period, December 1, 1976, to November 30, 1977, and this was the latest data available. A detailed description of a companion series of medical data can be found in Bosworth and Meyer [56].

We select ten fees as representative of those performed in the market. No attempt was made to systematically produce an index of procedures performed that was based on economic or dental theory. The following procedures were selected as representative of commonly performed procedures:

- Intraoral Film and Bitewings
- Intraoral Single First Film
- Intraoral Each Additional Film
- Prophylaxis Adults
- Amalgam One Surface Permanent
- Amalgam Two Surface Permanent
- Porcelain with Metal Crown
- Gold Full Cast Crown
- Complete Upper Denture
- Single Tooth Extraction

A fee index was constructed using the mix of services performed as relative weights. The index can be written as:

$$\text{Fee Index} = \sum_{i=1}^{10} W_i P_i$$

Where w_i is the number of procedures of type i that were performed divided by the total number of procedures of all types. This fee index has the desirable property of being easy to interpret as a weighted average of all procedures. However, it also has faults. One fault is that our structural equations are in logarithmic form while it can be shown that in this case an index constructed using the geometric mean rather than the arithmetic mean would improve the statistical precision of the standard errors. A second fault is that any possible changes in treatment are obscured by this method, since we are assuming that the mix does not change across regional markets. We have no reason to believe these two influences would substantially affect our results. However, they should be mentioned as potential problems, since this is the first study to employ the HIAA dental data base. The sixteen counties and county groups used in the study are listed in Appendix A. The remaining data sources are less unique and they will be mentioned only briefly.

The population data were obtained from Population of California Counties 1970-1976 published by the State Department of Finance, Population Research Unit. We used the total California population as of July 1, 1976. Personal income data came from the United States Department of Commerce, 1975. Median family income was published by the Department of Housing and Community Development for 1970. Our wage data was collected by the California State Department of Industrial Relations and published as Research Report Number 127 for the period January-March 1976. Population density in California was obtained from the State Department of Finance and was the estimate as of July 1, 1976. Population projections were also obtained from the State Department of Finance and the projection we used is the July 1, 1985 forecast. The percentage increase we used was the change in the July 1, 1976, to the July 1, 1985 period, divided by the July 1, 1976, base.

Medi-Cal Program payments to selected providers are reported in dollar amounts distributed to dental care recipients for the calendar year 1978. These data were provided by the California State Department of Health and are available in the *Medi-Cal Services and Expenditures Report*. The information on fluoride came from the United Fluoridation Census, United States Public Health Service.

Unionization data were provided by the California State Department of Labor and listed the number of members as of July 1977. The regional concentration of full time dentists was obtained from "The Distribution of Dentists," 1976, published by the American Dental Association.

Table 1

DEPENDENT VARIABLE	Constant	Fee	Population	Population/ Square Mile	Personal Income	Unions	Fluoride	Dentist/ Population	Wages	Fees x Procedures/ Population	%Δ Population Projection 1985	Savings & Loans	%Δ Population 1960-1970	Medi-cal
MODEL 1														
Demand (Procedures)	20.68	-6.00	2.36	-0.14	-0.45	0.59								-1.33
	(1.61)	(-3.03)	(2.64)	(-0.73)	(-1.64)	(.071)								(-2.11)
Supply (Procedures)	-44.72	6.40						3.23	8.47					
	(-2.06)	(2.74)						(2.40)	(3.66)					
Concentration (Dentist/ Population)	-9.95			-0.02	0.26	-0.70				0.17	-1.62	0.76	-0.01	
	-10.12)			(-0.36)	(2.68)	(-2.95)				(1.94)	(-0.04)	(2.64)	(-0.99)	
MODEL 2														
Demand (Procedures)	12.02	-3.48	1.94	-0.02	-1.10	-0.26								0.57
$r^2 = 0.93$	(0.87)	(-2.14)	(1.93)	(-0.09)	(-1.53)	(-0.87)								(0.60)
Supply (Procedures)	-42.31	4.39						2.86	8.71					
$r^2 = 0.76$	(-1.80)	(2.02)						(1.98)	(3.44)					
Concentration (Dentist/ Population)	-9.95			-0.02	0.26	-0.70				0.17	-1.62	0.76	-0.01	
$r^2 = 0.84$	(-8.92)			(-0.32)	(2.36)	(-2.60)				(1.71)	(-0.04)	(2.33)	(-0.87)	

Empirical Results

Since the data base for a number of the variables in the model has not been used before, it is appropriate to investigate a number of alternative specifications. This is more in the context of exploratory data anlysis, rather than testing modifications to well developed economic theory. The functional form for the equations in the model is double log, that is, natural logarithms of both dependent and independent variables. We had no reason to prefer one functional form to another and this specification allowed a straightforward interpretation of the coefficients as elasticities. As noted in the literature review no well developed theory of supply induced demand or locational concentration exists; thus our results are intended to be a first step in integrating standard economic analysis and some of the ad hoc research findings in dental practice location.

Model 1

The first model (I) displayed in Table 1 is the two stage least squares estimates of the coefficients with the t statistics in parenthesis. Our general impression is that some of the results from the model are in agreement with conventional economic theory. The price variable in the demand and supply equations have the expected signs as does population. The computed coefficients of income and unions do not have the expected sign but are not statistically different from zero. Medi-Cal expenditures presents a surprise. It was hypothesized that these expeditures would have shifted the demand curve for dental services to the right. Our empirical results seem to show that these governmental expenditures are in fact a proxy measure of low socio-economic status or poverty in the context of the demand model.

Another surprise was the positive sign for wages in the supply equation. It was expected that wages would be a measure of the cost of labor to the dental practice. In this context, an increase in costs would tend to reduce supply. If increases in regional income produce increases in dental fees, it may be that the quality of dental procedures are in fact a function of income. That is, procedures can have quality components not measured by the ADA procedure code. One group of unmeasured attributes could be the amenities consumed with the service itself. These amenities include the office and waiting room decor, the neighborhood where the office is located, the wait for an initial or subsequent appointments. Another component of dental service may be the time spent by the dentist and assistants in delivering the service. It might be that dentists (on average) supply "Chevrolet" services in lower income areas and "Buick" services in higher income areas.

Table 2

EPEN-ENT ARI-BLE	Con-stant	Fee	Popu-lation	Popu-lation/ Square Mile	Med-ian Income	Unions	Fluoride	Dentist/ Popu-lation	Wages	Fees x Proced-ures/ Popu-lation	%Δ Popu-lation Projec-tion 1985	Savings & Loans	%Δ Popu-lation 1960-1970	Medi-Cal
ODEL 3														
emand ro-	25.84	-4.14	2.33	-0.003	-0.91	0.90	-0.55							-1.90
edures)	(0.36)	(-1.22)	(1.96)	(-0.009)	(0.16)	(0.74)	(-1.23)							(-1.11)
ipply	-42.27	5.17						3.07	8.54					
edures)	(-1.84)	(2.31)						(2.16)	(3.49)					
oncen-ation entist/ opu-	-11.18			-0.13	0.43	-0.25				0.06	-72.33	0.43	0.02	
tion)	(-1.73)			(-1.62)	(0.59)	(-0.89)				(0.56)	(-2.10)	(1.07)	(2.06)	
ODEL 4														
emand ro- edures)	-4.16	-2.56	1.93	-0.10	1.42	0.62	-0.44							-1.23
= 0.94	(-0.07)	(-1.04)	(1.84)	(0.31)	(0.31)	(0.53)	(-1.03)							(-0.88)
ipply ro- edures	-42.31	4.39						2.86	8.71					
= 0.76	(-1.80)	(2.02)						(1.98)	(3.44)					
oncen-ation entist/ opu-tion)	-11.18			-0.13	0.43	-0.25				0.06	-72.33	0.43	0.01	
= 0.73	(-1.64)			(-1.54)	(0.56)	(-0.85)				(0.53)	(-2.00)	(1.02)	(1.96)	

nfortunately, we do not have any data on these atters and thus these statements are only conjectures this time.

The concentration equation seems to be in accordance ith our *a priori* expectations. Income of the potential atients and the general economic conditions of the area s measured by the number of Savings and Loan istitutions) are apparently major factors determining ie concentration of dentists. The proxy variable for surance has an opposite sign from our expectations. owever, in this equation it may also be a proxy easure for low socio-economic status of the area. The onstructed variables fees x procedures/population or er capita dental expenditures has the expected sign but s statistical significance does not meet customary andards. The introduction of past population growth id future population projections did not lead to a omplete answer to the question of whether oncentration leads or lags population growth. It opears that neither past growth nor future expected owth are influential determinants of dental oncentration ratios in this model.

LS Estimation of Model 1

Since we were estimating coefficients of over entified structural equations belonging to an interdependent system of equations, two stage least squares is an appropriate estimation technique. Using two stage least square leads to statistically consistent estimates of the parameters, and ordinary least squares (OLS) would not. However it may be interesting to see the OLS estimates of the structural equations.

The most striking contrast in the two models is the reduction in the absolute value of the price elasticity of demand and the price elasticity of supply. In the supply equation, the dentist population ratio is treated as exogenous and the computed value of the coefficient and t value falls. It appears that when the concentration ratio is assumed to be outside the interaction of market forces it has less importance in the supply of dental services. The location concentration ratio equation itself seems quite stable and robust to changes in estimation techniques.

Model 3 or 4 — Introduction of Fluoride Data

A source of data became available to us on the existence of fluoride in drinking water. This county data was in binary form, indicating the presence or absence of natural and artificial fluoridated water supplies. We assumed that the presence or absence of fluoride would not be a direct factor in either the supply or density equation.

Table 2 presents the two stage least squares estimates of Model 3. Comparing the computed estimates of this model with Model 1, the coefficient signs in the demand equation are the same. The computed coefficient for fluoride is negative but not significant, at least not at the 5 percent level. Note that median income was substituted for personal income as a measure of income less subject to relatively extreme values. The effect of introducing the binary fluoridation variable was to reduce the impact of income on demand for dental services.

Population as a separate influence in the demand equation becomes statistically questionable. The effect of population density on fee is reduced and the influence of unions (the proxy for insurance) is increased.

The supply equation in this model is much the same as in Model 1. It appears that the introduction of the zero-one variable for fluoridation has the effect of reducing the supply (quantity) response to changes in price. The remaining variables in the equation remain virtually unchanged. The dentist concentration ratio also appears to be robust to the introduction of fluoride to the model. All of the signs remain the same and the same general economic influences seem to prevail. However, the magnitude of the various exogeneous variables seems to have been altered. In particular, the percentage change in population between 1960 and 1970 change sign and the magnitude of the coefficient for forecasted population change greatly increased. This may be due to the change in the measure of income being altered from personal income to median income. Another explanation may be that an income measure which is less sensitive to extreme values increased the concentration equation's sensitivity to past and future population changes by removing the influence of those income fluctuations.

These results seem stable when we estimate the parameters by ordinary least squares. These coefficients are displayed in table two under the heading of Model 4. There, as in our previous comparisons, the effect of OLS estimation sheds some light on the effect that simultaneous estimation has on the computed parameter estimates. The simultaneity of the model appears to increase the measured influence that the dentist distribution equation has on dental service supply. In the supply equation, where the two stage least squares estimates account for the interdependence of supply and location, the distribution of dentists affects the supply of dental services. Where the dental distribution ratio is exogenous or predetermined and not part of the simultaneous nature of the model, dental distribution has much less statistical influence. The remaining variables have approximately the same measured relations as in Model 3 equation.

Dental demand is approximately the same under this estimation method as in the simultaneous model. Income becomes relatively more important and price becomes less important where the interaction of supply and demand are not specifically modeled. Again, population, unions, and fluoride have the expected signs. The sign of Medi-Cal is contrary to what we had expected. It may be that medical payments are in fact also measuring poverty or low socio-economic status or other income related phenomenon. It appears appropriate to estimate the model without this possible misspecification.

Models 5 and 6 — The Elimination of Medi-Cal Payments

Model 5 is Model 3 excluding the natural logarithm of Medi-Cal payments. The most striking feature here is the stability and robustness of the supply and concentration equations. The major changes were a slight drop in the coefficient for concentration indication, the possibility that the interaction of dental welfare payments via the dental services markets slightly reduced the economic forces determining the distribution of dentists. The concentration equation itself was virtually unchanged. It appears that, as one would expect, the influence of welfare related expenditures on the economic model is primarily through the demand equation. Supply of dental services is little changed by the variation in dental welfare payments. This may be due to the relatively small component Medi-Cal represents in total dental expenditures. Also, the Medi-Cal patients consumption mix of services may differ from those of the general population so dramatically that the inclusion of Medi-Cal in demand equation is suspect. The mix we used appears to be representative of what the typical or representative patient might consume. We do not know if this is an adequate measure for the dentally indigent population in California. Since the supply curve is relatively stable, it appears that our quantity index is appropriate for the combined population and that not much economic influence is exerted on the supply of dental services by dental welfare recipients.

Removing the Medi-Cal variable had the most noticeable effect on the parameter estimates for demand. The elasticity of demand increases and the income elasticity decreases. Although this is not conclusive proof, it appears that the demand specification is extremely sensitive to the measurement of income. The results of this equation are generally consistent with the research outlined above. The elasticity of demand appears to be low in this version of the model and the median income elasticity is high.

The ordinary least squares estimates are displayed in Table 3 as Model 6. In this case the demand elasticity is not appreciably altered by using OLS. In the other cases we found that the elasticities were reduced when the simultaneous nature of the system was modeled. Here we find the computed price elasticity increased, income elasticity decreased, and the remaining parameter estimates were about the same. Our estimate of the supply price elasticity falls in this model as it has in the OLS estimates. The dentist population ratio and cost

Table 3

DEPENDENT VARIABLE	Constant	Fee	Population	Population/ Square Mile	Median Income	Unions	Fluoride	Dentist/ Population	Wages	Fees x Procedures/ Population	%Δ Population Projection 1985	Savings & Loans	%Δ Population 1960-1970
MODEL 5													
Demand (Procedures)	-49.62 (-1.70)	-0.89 (-0.36)	1.53 (1.43)	-0.21 (-0.70)	4.73 (1.59)	-0.15 (-0.17)	-0.22 (-0.65)						
Supply (Procedures)	-47.65 (-1.99)	5.78 (2.43)						2.87 (2.01)	8.81 (3.61)				
Concentration (Dentist/ Population)	-11.18 (-2.20)			-0.13 (-2.07)	0.43 (0.75)	-0.25 (-1.14)				0.06 (0.71)	-72.33 (-2.69)	0.43 (1.37)	0.02 (2.63)
MODEL 6													
Demand (Procedures) $r^2 = 0.929$	-49.13 (-1.92)	-0.95 (-0.59)	1.55 (1.64)	-0.21 (-0.73)	4.68 (1.80)	-0.16 (-0.21)	-0.22 (-0.65)						
Supply (Procedures) $r^2 = 0.76$	-42.31 (-1.80)	4.39 (2.02)						2.86 (1.98)	8.71 (2.53)				
Concentration (Dentist/ Population) $r^2 = 0.73$	-11.18 (-1.64)			-0.13 (-1.54)	0.43 (0.56)	-0.25 (0.85)				0.06 (0.53)	-72.33 (-2.00)	0.43 (1.02)	0.02 (1.96)

coefficients are about the same. However, we find that the statistical significance of these estimates falls. Again, our finding is that the computed coefficients in the dentist concentration ratio are extremely robust to changes in estimation method.

Models 7 and 8 — Deletion of Los Angeles

One question came to mind during the evaluation of the research results. Since Los Angeles county is a major population center, does its presence in the data base drastically influence our results? We applied some of the techniques suggested by Tukey [54] and Mosteller and Tukey [55] and found no reason to reject the hypothesis that Los Angeles is a special case. We recomputed the parameter estimates of Model 5 omitting Los Angeles county and these results are displayed in Table 4 and labeled as Model 7.

In the demand equation the only meaningful change was to increase the computed price elasticity. However, the supply equation was altered. First we found a reduction in the supply price elasticity, an increase in the dentist density coefficient, and a decrease in the cost parameter. These results are consistent with market behavior in that demand appears to be stable across geographic areas after other economic influences are held constant. However, the presence or absence of Los Angeles County does affect supply opportunities. Apparently the presence of Los Angeles (perhaps via its geographic proximity to other population centers such as Orange County and San Diego County) increases the supply price response. Also, the dentist population ratio becomes more important, both in the magnitude of the coefficient and its statistical significance in this case. One interpretation is that the simultaneous nature of the market forces acting on the supply and location equations are most clearly measured when the Los Angeles market does not influence the estimates. Of course, we loose statistical efficiency by not including all the information we have on the sample. The only basis we would have for excluding the observation would be if we believed it were drawn from a different statistical population.

When we look at the concentration equation, we are again struck with the robustness of the computed coefficients. We are left with the impression that only supply is influenced by a change in the data base.

Model 8 contains the OLS estimates of the parameters of Model 7. The demand equation estimates and the

Table 4

DEPENDENT VARIABLE	Constant	Fee	Population	Population/ Square Mile	Median Income	Unions	Fluoride	Dentist/ Population	Wages	Fees x Procedures/ Population	%Δ Population Projection 1985	Savings & Loans	%Δ Population 1960-1970
MODEL 7													
Demand (Procedures)	-51.23 (-1.75)	-1.01 (-0.48)	1.56 (1.48)	-0.23 (-0.73)	4.94 (1.63)	-0.19 (-0.23)	-0.24 (-0.65)						
Supply (Procedures)	-25.58 (-0.99)	4.08 (1.72)						3.34 (2.40)	6.90 (2.59)				
Concentration (Dentist/ Population)	-11.61 (-1.70)			-0.14 (-1.57)	0.48 (0.62)	-0.25 (-0.89)				0.06 (0.56)	74.00 (2.11)	0.42 (1.08)	0.02 (1.70)
MODEL 8													
Demand (Procedures) $r^2 = 0.904$	-51.26 (-1.83)	-1.01 (-0.59)	1.56 (1.57)	-0.23 (-0.74)	4.95 (1.71)	-0.19 (-0.24)	-0.24 (-0.65)						
Supply (Procedures) $r^2 = 0.73$	-24.25 (-0.95)	3.34 (1.53)						3.19 (2.29)	6.90 (2.57)				
Concentration (Dentist/ Population) $r^2 = 0.73$	-11.61 (-1.49)			-0.13 (-1.73)	0.48 (0.54)	-0.25 (-0.78)				0.06 (0.49)	-74.00 (-1.85)	0.42 (0.94)	0.02 (1.49)

density equation estimates are surprisingly close to the two stage least squares estimates. The supply price elasticity falls using OLS estimation as does the dentist location coefficient, and these results are consistent with the results in our other models.

Model 9 and 10 — Supply Induced Demand

A final run was made including the concentration ratio in the demand equation and those results are presented in Table 5 labeled Model 9. We attempted to see if "supply induced" demand would be confirmed in our three equation model.

In this case, it is clear that we should consider the correlation of the disturbances across equations. If we do not consider this correlation in the three structural equations we will have estimates which are not asymptotically efficient. In effect we have prior restrictions across equations which are not included in our model. Rather than the single equation technique of two stage least squares (which in the context of our previous models were statistically consistent) which was satisfactory in our other models, here we need a system approach such as full information maximum likelihood. The estimates presented for Model 9 were computed using the full information maximum likelihood technique. We found a positive statistical relation between the amount of services performed and the concentration ratio. This would apparently confirm the supply induced demand hypothesis. However, when the whole equation is examined we find a different result. The coefficient for price is now positive and the income elasticity falls sharply. The computed signs for population density and unions are reversed. Fluoride continues to exert the expected negative impact on quantity of dental services demanded.

Apparently dental concentration and population exert almost all of the statistical effect on the quantity of procedures demanded. This statistical relation does not resemble a demand function. The fee for service has a computed sign contrary to fundamental economic analysis. Also income exerts a very small influence on consumption — a finding at odds with other work.

The supply and concentration equations are also inconsistent with other findings. The supply elasticity is unreasonably large. The coefficient for dental concentration increases greatly and is only understandable after looking at the demand equation. The dentist distribution equation seems to be driven by

Table 5

DEPENDENT VARIABLE	Constant	Fee	Population	Population/ Square Mile	Median Income	Unions	Fluoride	Dentist/ Population	Wages	Fees x Procedures/ Population	%Δ Population Projection 1985	Savings & Loans	%Δ Population 1960-1970
MODEL 9													
Demand (Procedures)	-9.43	2.04	0.88	0.04	0.62	0.14	-0.58	3.25					
	(-5.33)	(1.82)	(+100.00)	(0.24)	(0.42)	(1.08)	(-1.45)	(8.08)					
Supply (Procedures)	-42.79	24.44						9.31	0.002				
	(-2.89)	(2.60)						(2.76)	(2.64)				
Concentration (Dentist/ Population)	1.07			-0.02	-0.29	-0.65				0.41	-.001	.0005	.0000000
	(10.25)			(-.24)	(-.42)	(-1.08)				(+100.00)	(-1.37)	(2.27)	(0.75)
MODEL 10													
Demand (Procedures)	-4.64	-2.28	2.15	0.08	-3.57	1.18	-0.41	2.23					
$r^2 = 0.93$	(-2.07)	(-1.27)	(2.78)	([illegible].34)	(-0.95)	(1.35)	(-1.32)	(2.81)					
Supply (Procedures)	-42.31	4.39						2.86	8.71				
$r^2 = 0.76$	(-1.80)	(2.02)						(1.98)	(2.53)				
Concentration (Dentist/ Population)	-11.18			-0.13	0.43	-0.25				0.06	-72.33	0.43	0.02
$r^2 = 0.73$	(-1.64)			(-1.54)	(0.56)	(0.85)				(0.53)	(-2.00)	(1.02)	(1.96)

per capita expenditure and to a lesser extent general economic conditions as measured by the number of savings and loan institutions. This result is consistent with the hypothesis that market forces generate the distribution of dentists. That hypothesis appears to be extremely robust to alternative market formulations. However, the results from the demand and supply equations lend little support to the supply induced demand hypothesis.

As we did in the other models, we present the OLS estimates of the parameters of the demand equation. The other equations would, of course, remain the same, and they are presented in Table 4 for comparison. In this we see that the fee elasticity has the expected negative sign. However, the computed income elasticity is negative. Again it appears that population and dentist per capita exert much of the explanatory power in the equation. The results from this equation shows as dramatically as any others the need to use the appropriate statistical methods when estimating simultaneous systems of economic relations. The ordinary least squares estimates yield plausible results for some of the parameters. Other parameter estimates are less plausible. Fortunately, we have the full information maximum likelihood estimates, which appear to be the proper ones in the context of the model.

Summary

Our most important finding is that dental concentration appears to be a result of market forces. Our model combines orthodox economic theory of supply and demand for dental services with a third equation determining dental concentration. These three equations form a simultaneous system of economic activity. In traditional economics we know that under general circumstances we cannot estimate the parameters of either the demand or supply equation alone. We must consider the interaction between supply and demand. In our study we have some evidence that the distribution of dentists is also the result of interactions of supply and demand.

The policy implication of this finding is important. Since the market concentration of dentists or what has been called the "maldistribution" of dentists is a result of market forces, we must be particularly careful to fully understand the consequences of policies designed to alter the geographic distribution of dentists. This

could be put another way. If we intend as a public policy to greatly alter the geographic distribution of dentists, we will have to make changes that have great effects on the markets for dental services. This finding is consistent with the facts that dentist "maldistribution" is not unique to the United States. In our literature review we found that across a wide spectrum international health institutions fewer dentists per capita were associated with rural low income areas. Our findings are consistent with the view that even in an economically prosperous region — California — the same results occur. It seems that the fundamental forces that determine the demand and supply of dental services also simultaneously determine the distribution of dentists.

Another important finding is that knowledge can be gained by investigating smaller geographic areas than whole states. We would expect that counties or county regions are an important area of analysis. Future studies will of course refine this idea, but it seems clear that the next step is to test this model within other states. One interesting advancement would be to extend this model across states with differing licensing laws, practice acts, and regulation of auxiliaries. We could then see the influence of those characteristics that differed from state to state on supply, demand, and concentration at the same time measure the effect of competitive market forces within county areas.

Our estimates of the supply and demand equations were not as satisfactory as we had hoped. It is true that we do not have many direct comparisons where data as micro as ours were employed. Nevertheless, our parameter estimates are generally consistent with economic theory. Our estimates of unitary elasticity for dental demand are consistent with other findings. Our income elasticity estimate is large. The most probable area to improve our results would be to investigate the measure of income. We found that the measure of income made an important difference in our parameter estimates. It may be that we need to consider other characteristics of the income distribution in addtion to the mean or median. Measures of income dispersion, levels of poverty, or distribution of income at various levels could be important, since the relation between income level and dental service consumption may be complex. The complexity may only be evident when a fairly sophisticated model is employed and totally lost when fees are simply regressed against exogenous variables.

The supply equation is generally satisfactory, except that we are apparently measuring income rather than costs with the proxy variable of wages. One interpretation of our results for this equation is that the supply of services is in fact influenced by the distribution of dentists and in turn influences the location distribution via the per capita total revenue variable in the location equation. An increase in wages, which apparently act as a proxy for income, shifts the supply of dental services to the right. This may be due to a change in the mix of services performed or that the unit of service changes for higher income patients. This result would occur in the case where the listed or classified procedures were not unique in an economic sense. In this case, the quality of the dental procedure or service would vary directly with the income of the patient and the particular service actually being performed would not be the same for all patients. We would need to model a number of service qualities and amenities such as procedure time, waiting time for first appointment, and waiting time in office, to account for these quality influences.

The dentists concentration or distribution equation is remarkably stable and robust to alternative extimation techniques and alternative specifications. One interesting finding is the confirmation that the concentration of dentists does not anticipate population trends. We would expect that higher levels of dentist concentration would exist in areas of higher expected population growth. We do not find this. We find that the establishment of dental practices seems to follow past growth trends. The positive relation between the number of savings and loan institutions and dentist density must be a measure of the after tax discretionary income of the potential patient pool. This discretionary income would be related to current income, permanent income, and the age distribution in the area. In our results we found median income of the area to have positive relation to dentist concentration. These last two results are consistent with other findings that dentists seem to prefer areas with many social amenities and good schools, since these products are usually more readily available in higher as compared to lower income areas.

We found little evidence to support the supply induced demand hypothesis. When the concentration ratio of dentists was introduced into the demand equation a positive relation was found. However, that change to our model apparently generated a severe specification error. The result was a statistical relation that was not in fact a market demand function. Our model seems to show that the supply induced demand phenomenon occurs in cases where the full simultaneous nature of dental markets is not recognized.

References

1. Fuchs VR, Kramer, MJ: Determinants of Expenditures for Physicians' Services in the United States, 1948-1968. National Bureau of Economic Research, Occasional Paper #117, December 1972.
2. Feldstein MS: The Rising Price of Physicians' Services. *The Review of Economics and Statistics* LII: 121-133, 1970.
3. Arrow, KJ, Capron, WM: Dynamic Shortages and Price Rises: The Engineer-Scientist Case. *Quarterly Journal of Economics,* May 1959.

4. Rayack, E: The Supply of Physicians' Services. *Industrial and Labor Relations Review,* January 1964.
5. Fein, R: The Doctor Shortage: An Economic Diagnosis. *The Brookings Institution,* Washington, 1967.
6. Newhouse, JP: A Model of Physician Pricing. *Southern Economic Journal* 37:17-183, 1970.
7. Evans, RGL: Supplier-Induced Demand: Some Empirical Evidence and Implications. *The Economics of Health and Medical Care.* (Pereman, M., Ed.). London, Macmillan Co., 162-173, 1970.
8. Evans, RG, Parrish, EMA, Sully, F: Medical Productivity, Scale Effects and Demand Generation. *Canadian Journal of Economics* VI, 1973.
9. Frech, HE, Ginsburg, PB: Physician Pricing: Monopolistic or Competitive: Comment. *Southern Economic Journal* 38:573-577, 1972.
10. Kessel, R: Price Discrimination in Medicine. *Journal of Law and Economics.* 1:20-53, 1958.
11. Hyde, DR, Wolff, P.: The American Medical Association: Power, Purpose and Politics in Organized Medicine. *Yale Law Journal* 63:943-975, 1954.
12. Friedman, M, Kuznets, S: Income From Independent Professional Practice. National Bureau of Economic Research, 1945.
13. Dyckman, Z: A Study of Physicians' Fees. Staff Report Council on Wage and Price Stability. USGPO, Washington, March, 1978.
14. Brown, DM, Lapan, HE: Rising Price of Physician's Services: A Comment. *Review of Economics and Statistics* LIV:101-105, 1972.
15. Feldstein, M: The Rising Price of Physicians' Services: A Reply. *Review of Economics and Statistics.* LIV: 105-107, 1972.
16. Freidson, E: Profession of Medicine. Dodd, Mead and Co., New York, 1971.
17. Martin, W: Preferences for Types of Patients. The Student Physicians by Merton, K.R., et al., Harvard University Press, Cambridge, 1957.
18. Reder, MW: Some Problems in the Measurement of Productivity in Medical Care Industry. Production and Productivity in the Service Industries. Columbia University Press, New York, 1969.
19. Steinwald, B, Sloan, F: Determinants of Physicians' Fees. *Journal of Business.* 47:493-511, 1974.
20. Kushman, JE, Scheffler, RM: Pricing Health Services: Verification of the Monopoly Pricing Model for Dentists. *Journal of Human Resources* XIII: 402-415, 1978.
21. Newhouse, PE, Sloan, FA: Physician Pricing: Monopolistic or Competitive: A Reply. *Southern Economic Journal* 39:258-262, 1972.
22. Holtman, AG, Olsen, EO: The Demand for Dental Care: A Study of Consumption and Household Production. *Journal of Human Resources* XI:546-560, 1976.
23. Shepard, L: Licensure Restrictions and the Cost of Dental Care: An Empirical Analysis. University of California, 1976.
24. Maurizi, AR: Public Policy and Dental Care Market. American Enterprise Institute, Washington, D.C., 1975.
25. Acton, JP: Non-monetary Factors in the Demand for Medical Services: Some Empirical Evidence. *Journal of Political Economy* 83:595-614, 1975.
26. Sloan, FA, Lorant, JH: The Allocation of Physicians' Services: Evidence of Length of Visit. *Quarterly Review of Economics and Business* 16:85-103, 1976.
27. Sloan, FA, Lorant, JH: The Role of Waiting Time: Evidence from Physicians' Practices. *Journal of Business* 50:486-507, 1977.
28. Pashigian, BP: Occupational Licensing and the Interstate Mobility of Professionals. National Bureau of Economic Research, Mimeo, February, 1978.
29. Glass, RL, Baldwin, DC: Factors Influencing the Selection of Dental Office Location. *Journal of the American Dental Association* 1337-1341, December, 1970.
30. Wash, P: Occupational Mobility of Health Workers. *Monthly Labor Review* 100:25-29, 1977.
31. Yett, DE, Sloan, FA: Migration Patterns of Recent Medical School Graduates. *Inquiry* 11:125-142, 1974.
32. Weiskotten, HG, et al. Trends in Medical Practice: An Analysis of the Distribution and Characteristics of Medical College Graduation, 1915-1950. *Journal of Medical Education* 35:1071-1121, 1960.
33. Benhan, L, Maurizi, AR, Reder, MW: Migration, Location and Remuneration of Medical Personnel: Physicians and Dentists. *Review of Economics and Statistics* L:332-347, 1968.
34. Tryon, AF: Determinants of Variations in the Supply and Distribution of Dental Manpower Among Metropolitan Areas. *Journal of Dental Research* 52:774-778, 1973.
35. Tryon, AF: Factors Influencing Variations in Distribution of Dental Manpower in an Urban Area. *Public Health Reports* 89:320-324, 1974.
36. Rimlinger, GV, Steele, HB: An Economic Interpretation of the Special Distribution of Physicians in the U.S. *Southern Economic Journal* XXX:1-12, 1963.
37. Steele, HR, Rimlinger, GV: Income Opportunities and Physician Location Trends in the United States. *Western Economic Journal* III:182-193, 1965.
38. Ball, DS, Wilson, JW: Community Health Facilities and Services: The Manpower Dimensions. *American Journal of Agricultural Economics* 50:1208-1222, 1968.
39. American Dental Association: Survey of Recent Dental Graduates: Factors Related to Selection of a Practice Location. *Journal of the American Dental Association* 87:904-906, 1973.
40. Robert Wood Johnson Foundation: Bibliography of Physician Location. Human Resources Research Center, Robert Wood Johnson Foundation. Publication Number 300, 1974.
41. Dean, RT, McClung, RL, Tobin, SJ: Economic Factors in Dentist Maldistribution. Applied Manage-

ment Sciences. Silver Spring, Maryland, October, 1977.

42. Armstrong, RW: Medical Geography — An Emerging Specialty? *International Pathology* 6:61-63, 1965.
43. Dever, AG: Medical Geography — A Selected Bibliography. *International Pathology* II:13-16, 1970.
44. May, JM: Medical Geography: Methods and Objectives. *Geographical Review* 40:9-41, 1950.
45. Shannon, GW, Denver, GE: Health Care Delivery: Spatial Perspectives. McGraw Hill, 1974.
46. Elesh, D, Schollander, PT: Race and Urban Medicine: Factors Affecting the Distribution of Physicians in Chicago. *Journal of Health and Social Behavior* 13:236-250, 1972.
47. Guzier, DS, Jahiel, RI: Distribution of Private Practicing Offices of Physicians with Specific Characteristics Among Urban Neighborhoods.. *Medical Care* XIV:469-488, 1976.
48. Duff, RS, Hollingshead, AB: Sickness and Society. Harper and Row, New York, 1968.
49. Markovin, AP: Historical Sketch of the Development of Soviet Medical Geography. *Soviet Geography* 3:3-19, 1962.
50. Roemer, MI: Strategies for Increasing Rural Manpower in Five Industrialized Countries. *Public Health Reports* 93:142-146, 1978.
51. Holen, A: The Economics of Dental Licensing. Monograph. The Public Research Institute, Center for Naval Analyses, Mimeo, October, 1977.
52. MacBride, OE: Restrictive Licensing of Dental Para-Professionals. *Yale Law Journal* 83:806-826, 1974.
53. Sheppard, HJ: Testimony Presented to the Council of State Governments', National Task Force on State Dental Policies, (Available Through Federal Trade Commission, San Francisco Office), October 24, 1978.
54. Tukey, JW: Exploratory Data Analysis, Addison-Wesley Publishing Company (Reading, Mass.), 1977.
55. Mosteller, F, Tukey, JW: Data Analysis and Regression, Addison-Wesley Publishing Company Reading, Mass.)., 1977.
56. Bosworth, BP, Meyer, J: A Study of Physicians' Fees, Council on Wage and Price Stability, Executive Office of the President, March, 1978, especially Chapter Five, pp. 95-126 and Appendix 2, pp. 137-140.
57. Upton, C, Silverman, W: The Demand for Dental Services, *Journal of Human Resources*, VII:250-261, 1972.
58. Feldstein, P: Financing Dental Care: An Economic Analysis. Lexington Books, Massachusetts, 1973.

APPENDIX A

Health Insurance Association Zip Code Dental Areas

Zip Codes	*California County Names*
956, 958	Sacramento, Yolo, Placer, El Dorado, Amador
935, 936, 937	Fresno, Madera, Inyo, Mono
932, 933	Kern, Tulare, Kings
935	Merced, Mariposa, Stanislaus, Tuolumne
939	Monterey
926, 927, 928	Orange
922	Riverside, Imperial
920, 921	San Diego
952	San Joaquin, Calaveras
931	Santa Barbara
950, 951	Santa Clara, Santa Cruz, San Benito
954	Sonoma, Mendocino, Lake
930, 034	Ventura, San Luis Obispo
900-918	Los Angeles
923, 924, 925	San Bernardino
940, 941, 945-949	San Francisco, San Mateo, Marin, Alameda, Contra Costa, Solano, Napa

Conflicting Theories of the Determination of the Entrepreneur's Income: An Analysis of the Practicing Dentist

Donald R. House, Ph.D.
American Dental Association

ABSTRACT □ Current published research relies principally on the association between the provider to population ratio and fee levels as "proof" of a fundamental failure in the health care market mechanism. This association appears to be the basis for interest in the target income hypothesis.

Since the literature presents only sketchy notions of targeted income behavior, this paper examines possible theoretical constructs of provider behavior. Both constrained and unconstrained demand creation are examined in the context of profit maximization and targeted income objective functions. Among the several constructs of a targeted income model examined in the paper, we are unable to record any substantive contribution from such deviations from the traditional market-oriented models.

An alternative formulation of a dental care market is explored wherein patients must expend both time and money for the purchase of the dentist's services. Under quite plausible conditions in a competitive market, we find that the observed association between firm density and fees is a probable result, consistent with allocations in an efficient market. Indeed, the competitive model with consumer search and stochastic patient arrivals yields such results.

Since the early 1960's, when government's participation in health care markets began its most rapid growth, society has expressed a new interest in the incomes of physicians and dentists. Rapid inflation in these markets spurred Congressional consideration of fee and income determination of health professionals. Legislation that expanded the number of physicians, dentists, and hospitals did not seem to retard the growth of prices of health care. Many have argued that health care is a "different" commodity or services for which conventional economic tools of analysis are inappropriate.

Several of the earlier papers published in economic journals offered some unexpected empirical results. Cross-sectional studies led to the conclusion that fees and firm density are directly related (see Fuchs-Kramer [7], Kushman-Scheffler [11], and Newhouse [12]). National time-series data were used to question the existence of a market-clearing mechanism in the market for physician's services (see Feldstein [6]). In response to these preliminary findings, new notions of physician and dentist behavior were advanced, which has led to recent considerations of the target income hypothesis. Presently, no one has yet developed a fully formalized theory of targeted behavior with considerations of market allocations and clearing mechanisms. Without such a theory, it is impossible to evaluate the merits of this new approach to health care markets.

This paper offers an examination of the theoretical substance of two related notions of health provider behavior — demand creation and the target income hypothesis. Hopefully, this examination will clarify the meaning of these concepts and offer some direction as to how they might be consistently introduced into a theoretical framework. While it promises no definitive test of the target income hypothesis, this paper offers a general evaluation of its merits as a viable alternative to conventional theory of firm behavior. We find that

evidence of the direct association between firm density and price is clearly consistent with a more traditional approach.

DEMAND CREATION AND TARGETED INCOMES — AN EXAMINATION OF MEANING

Perhaps the most widely accepted alternative notion of provider behavior is the demand creation phenomenon. It appears often in empirical sections of papers wherein variations in consumer demand are adjusted for the direct effect of provider influence on patient decisions, often termed "supplier-induced demand" (see, for example, Fuchs-Kramer [7]). Although many demand-estimating equations include the provider/population ratio as an adjustment, the arguments on behalf of its inclusion differ; the theoretical basis for its inclusion lies at the heart of the demand creation issue.[1]

On the one hand, the conventional demand-supply model with recognition of time prices leads to a relevant consideration of firm density in demand estimations. With the introduction of patient travel time and patient queues, money prices or fees no longer represent the marginal expense of services consumed.[2] Models which incorporate patient time normally regard time as simply a component of full price. The usual separation between the provider's and consumer's decision-making processes is maintained. That is, each participant makes decisions based upon market-determined variables (money and time prices), and direct influences upon each other's activities (such as supply inducement) are either absent or so small that they can be ignored.

On the other hand, inclusion of the provider/population ratio is sometimes based upon an assumed *direct* influence of provider over consumer. The provider in some way manipulates demand through influence upon the patient's decision-making process (some view the *provider* as the relevant decision maker). Patient ignorance of physiological functions and reactions to the chemical compositions of medicines often foster the presumption of this theoretical link between the provider and his patient's demand, independent of money and time prices.[3]

With the alleged control over patient decisions, estimation of demand functions should incorporate a measure of the provider's direct influence. This measurement of direct influence, according to many, is mysteriously captured in the provider/population ratio. Without formal consideration of firm behavior, these empirical studies implicitly argue that the provider's desire to create demand is determined in part by this ratio. Increases in the ratio ration patients over a larger number of providers, and, in response, providers create

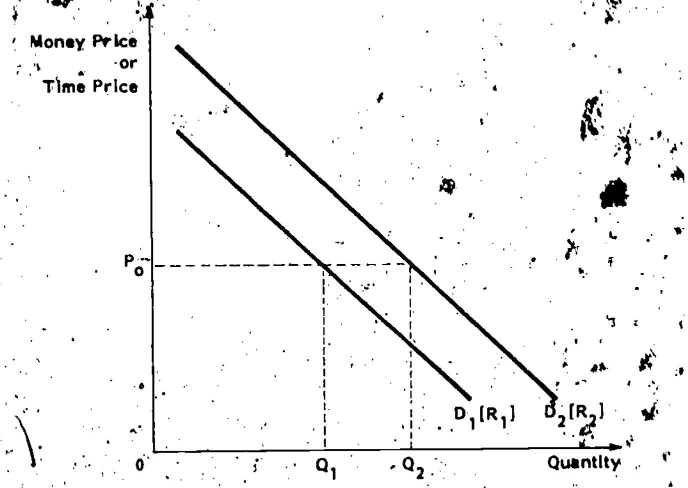


Figure 1

demand so as to maintain the same level of service delivery.

With this approach to market performance, one can expect a "perverse" response to firm entry. As more providers enter a market, increasing the provider/population ratio, demand will increase and *prices could rise*. Some have regarded the "perverse" association between the provider/population ratio and fees as "proof" of the failure of demand supply models to explain observed market changes.[4] A later section of this paper illustrates that a direct relation between the provider/population ratio is not at all perverse in cross-section studies. Such empirical associations cannot be regarded as a basis for rejection of conventional models.

An Operational Statement of Demand Creation

The literature does not go far in clarifying direct demand creation. (For our purposes, "demand creation" refers only to *direct* provider influence, not consumer reaction to prices.) Consider demand creation as an activity that can be varied according to the objectives of the provider. An abstract measure of this creation activity, although unobservable, is assumed to exist. If the variable R measures this creation activity, it is appropriate to include R in the patient demand function. Note that we do not eliminate the patient's response to money and time prices for a *given level of creation activity*. What we must consider are the economic determinants of R and the optimal level of demand creation activity by the provider.

Graphically, we can justifiably construct the consumer demand curve, as in Figure 1. The initial demand curve is presented as D_1 which reflects a level of demand creation, R_1. As the provider chooses a greater level of creation activity, R_2, the demand curve shifts to the right to D_2. This shift occurs in both price-quantity

---

[1] A comparison of explanations found in Fuchs-Kramer [17] and Newhouse-Phelps [13] offers a useful review of rationalization for inclusion of the provider/population ratio in demand functions.

[2] Acton [1] has clearly recorded the importance of time prices for medical services, and Holtmann-Olsen [9] did the same for dental care.

[3] Newhouse [12] underscores the importance of consumer ignorance in both physician and dentist markets. Price serves a very limited allocation role since patients choose firms largely on the basis of queues.

[4] See, for example, Evans [4], Newhouse [12], Scheffler-Kushman [16], Fuchs-Kramer [7], and Kushman-Scheffler [11].

and time-quantity spaces as it is expected that money prices and time prices respond to changes in R.

The mechanics of this demand creation appear at this stage as a simple advertising model. The variable R measures real advertising expenditures that influence consumer purchases. A basic departure from such models, however, lies in the possibility that demand creation in health care markets may be virtually *costless* because of the alleged provider control over patients. That is, the dentist supposedly can schedule an increased list of treatment visits with no additional expense (independent of any costs of delivery of treatment). If true, the demand curve facing the firm can be increased at will. The entrepreneur can choose any output that is "comfortable" and establish virtually any fee. The firm's output and price are chosen so as to provide optimal levels of labor time requirements and a desired income. While this set of assumptions runs counter to those of most economists because of the total disregard for market constraints, a quick review of the literature will uncover the conclusions in Dyckman [3], which represent the current administration's view of the surgeon's market. This finding alone gives immediate importance to the concept, and it is perhaps best to continue with a thorough consideration of implications for empirical testing.

Costless Demand Creation — A Basis for Targeted Income Behavior

Consider a model in which demand creation is virtually free; that is, the provider can manipulate patients and thereby shift the demand curve outward without money or psychic cost. Clearly, if the provider is a profit maximizer, any model incorporating these constructs will yield explosive solutions. Demand will shift outward, *ad infinitum*. Without market constraints, the provider need not produce health care with a least-cost combination of inputs, choose an output that equates marginal revenue and marginal cost, or supply more than a modicum of his own labor time. Casual evidence of the physician's long work week leads to a quick rejection of this theoretical approach. Its testable implications that are disproved by published evidence are too numerous to list. If costless demand creation is to remain a promising assumption in an empirically supported theory of firm behavior, we must introduce a logical limit — a constraint — to the provider's income-generating protential.

In search of a logical limit to demand creation activities, the notion of a targeted income was introduced.[5] Not to be confused with income expectations or projections, the target income hypothesis argues that there is an income limit that the provider chooses not to surpass for one reason or another. What follows is a consideration of possible alternatives that might be considered in search of a binding constraint to income generation.

Perhaps the simplest approach to targeted income behavior is a mere assumption that a target exists which in itself is the binding constraint. Evans has argued that we need not understand the determination of this target. However, if we adopt the assumed costless demand creation, targets will always be achieved. Being unconcerned about target determination is equivalent to a disregard for the income distribution among providers. The targeted income behavior has been used as a logical basis for limiting the supply of future dentists and physicians. Each new provider will undoubtedly claim health care expenditures in proportion to his targeted income. For this reason, it seems that even those following Evans would consider target determination in developing an effective cost containment policy. We need only select medical and dental students with relatively low targets!

Some published results offer some insight into consideration of target determination. Recently, Shepard [17] presented an income determination equation for dentists. His evidence suggests that dentists' incomes vary in predictable directions with changes in relevant market conditions. If we adopt costless demand creation, we are forced to accept the notion that targets vary predictably with market conditions. This becomes an empty hypothesis of provider behavior. Dentists *can* earn anything they want because of demand creation opportunities. Distributions of dentists' incomes must then reflect differences in tastes or preferences, not market conditions. It must be mere coincidence that dentists prefer no more and no less than what could be earned in a constrained market.

As an alternative specification of target determination, Feldstein [6] suggests that physicians may respond to a "reference income," that is, physicians (and dentists) wish for an income commensurate with that of someone in the community income distribution. If the community's income improves, the dentist will improve his preferred target. Presumably, the rationale for this reference income determination of targets alludes to professional guilt, risk of government regulation, or interdependent utility functions. While this offers some basis for explaining rising fees in areas with larger provider/population ratios, it misses the mark when it comes to explaining physician and dentist migration. With costless demand creation, migration patterns can be explained only with non-economic factors. Migration studies such as Benham-Maurizi-Reder [2] serve as a basis for rejection of this reference income determination of targets when coupled with costless demand creation.

In summary, costless demand creation seems to be useful in explaining only a single aspect of market performance: price response to entry of additional providers in a local market. Logically, there must be

[5] It is not clear who must be credited with the first serious consideration of targeted income behavior. Feldstein [6] perhaps should be credited for the notion of a functionless price mechanism while Fuchs-Kramer [7] argued on behalf of demand creation. The most spirited disciple of the target income hypothesis is clearly Robert G. Evans as noted in Evans [4] and Evans, Parish, Sully [5].

some binding constraint to demand creation, and the targeted income appears to be a popular one. Yet the literature does not address target determination nor reconcile implications with existing evidence that relates actual incomes to economic variables. In short, costless demand creation with targeted incomes successfully explains a "perverse" relation between fees and provider/population ratios. But it fails to explain other empirical evidence. With this, one cannot conclude that a preferred theory has been uncovered.

Demand Creation, Expenses, and Targeted Income Behavior

The introduction of some direct costs associated with demand creation offers a somewhat more promising approach to provider behavior. The dentist or physician is capable of creating demand but such creation is not costless. Again, some targeted income exists which the provider chooses not to surpass. This particular theoretical constraint offers unambiguous solutions at times, but actually is equivalent to existing models of firm behavior.

Consider a dentist with a clientele consisting of members who will not switch dentists. If fees are increased, however, some will choose to forego treatment. Hence, the demand curve facing the firm is negatively sloped similar to the traditional monopolist. Demand creation can be described as the outcome of solicited examination among members of a clientele. A reminder phone call to an existing member will usually bring the patient in for an examination. This provides the opportunity to add members to the queue waiting for treatment. Thus, demand has been "created."[6]

Let the dentist's utility function be represented by the equation:

$$U = U(E,H) \quad \frac{\partial U}{\partial E} < 0 \text{ and } \frac{\partial U}{\partial H} > 0 \qquad (1)$$

where $E = |y - y_t|$ and H represents household-produced services for which $H = H(\ell)$ where ℓ represents the dentist's chosen household time (or time away from the dental practice). y represents the net income of the practice (net of expenses associated with production of services and creation of demand). y_t represents a target income, the determination of which is unknown.

The demand function facing the firm relates price to both output X, and the dentist's demand creation activities, R. Hence: $P = P(X,R)$ where P represents the firm's price of the output (or the dentist's fee). The dentist's net income function is thus:

$$y = P(X,R) \cdot X(L,R) - P_k K - R \qquad (2)$$

where L represents the dentist's own labor time, K represents rented physical capital, P_k represents the rental price, and R measures the expenses associated with demand creation.

[6]Alternatively, demand creation could be defined in terms of delivering period oral exams below cost where R represents the total difference between revenues and costs associated with that activitiy.

The dentist faces a simple time constraint ($1 = \ell + L$) which effectively limits his choices of household time-labor time combinations. Combining these equations into a single Lagrangian function, we translate the concept of utility maximization with a target income into a mathematical maximization process.

$$Z = U[E(y,y_t), H(\ell)] + \lambda\{P[R,X(K,L)] \cdot X(K,L) - P_k K - R - y\} + \sigma\{1 - \ell - L\} \qquad (3)$$

The maximization of Z, the Lagrangian function, requires satisfying the following first order conditions:

$$\text{a)}\ \frac{\partial U}{\partial E}\frac{\partial E}{\partial y}\left[\frac{\partial P}{\partial R}X - 1\right] = \frac{\partial U}{\partial E}\frac{\partial E}{\partial y}\left[\frac{\partial X}{\partial K}P\left(1 - \frac{1}{\eta}\right) - P_k\right] = 0 \qquad (4)$$

$$\text{b)}\ \frac{\partial U}{\partial E}\frac{\partial E}{\partial y}\left[\frac{\partial X}{\partial L}P\left(1 - \frac{1}{\eta}\right)\right] = \frac{\partial U}{\partial H}\frac{\partial H}{\partial \ell} \qquad (4)$$

where η is the price elasticity of the demand curve facing the firm.

The key to interpreting these conditions lies in the sign of the term $\frac{\partial E}{\partial y}$. If the dentist has established a target income that exceeds the optimal income y, $\frac{\partial E}{\partial y} < 0$ and the first order conditions reduce to those of a conventional model. Demand creation activities, R, will be employed until its marginal return, $\frac{\partial P}{\partial R}X$, equals its marginal cost. Capital is employed until its marginal revenue product equals its rental price, and the dentist will supply labor time to the firm until the value of the incremental income equals its opportunity cost: the value of household time. It is important to note that these results are exactly identical to those of conventional models that include some form of advertising or selling costs for the product. The entrepreneur will maximize profits, *given* his choice of household time, by optimally employing advertising services.

Next, consider a dentist who is fortunate enough to have adopted a sufficiently small income target. While this particular specification might appear to offer a substantive break from conventional models, we find that, in fact, it offers nothing new.

If the dentist is capable of attaining his target without great difficulty, the existence of the household sector ensures stable, unique solutions in the firm. For every income increment the dentist extracts from the firm by combining capital and demand creation activities in a profit-maximizing manner, he "purchases" more time away from the practice. Even though his target is maintained, his firm is efficiently managed. Given a level of houshold time, the dentist attempts to maximize profits. Or, equivalently, given his targeted profits, the dentist attempts to minimize his own labor time input.

In the former case where the target clearly exceeds the optimal income level, we find that the labor-leisure

analysis is exactly equivalent to conventional analysis. If targets are never achieved, the model reduces to the utility-maximization-profit-maximization model that is long since been fully developed in the literature. In such models, an exogenous increase in demand (perhaps caused by a reduction in the dentist/population ratio) will lead to an increase in price and an increase in the marginal return to the dentist's own labor time. Whether he increases his labor time or not depends upon the domination of the income effect produced (or the backward bending supply curve of labor). Perhaps, at some point, the dentist will choose less labor time, but conventional theory argues that *income will never fall.* (This is essentially an extension of the argument posited by Lionel Robbins in the 1930's: the demand for income in terms of effort is never positively sloped.)[7]

In the latter case, an income level exists that will not be exceeded. That is, at some point (where the target is attained) the supply of labor curve will bend backward with an elasticity equal to unity. Income will never fall as labor time decreases in response to increasing demand for dental care.

In comparison to existing theory, the two specifications of targeted levels offer no substantive contribution. The analysis is essentially identical with the exception that the targeted income approach places a greater restriction upon the elasticity of the supply of labor curve at the dentist's wage where the targeted income is attained. In both instances, dental firms are managed as profit-maximizing entities, given the dentist's labor supply choice. Demand creation as viewed as a patient screening activity is utilized in a manner that maximizes profits even though targeted income behavior is operative. As presented, the target income hypothesis offers a virtually identical theoretical basis for explaining firm behavior. Presumably, the only deviation in predictions is a subtle restriction concerning the elasticity of the labor supply curve as some wage that is unknown and unpredictable.

In summary, the demand creation model incorporating targeted income behavior and direct creation expenses seemingly offers no substantive improvement as a theoretical basis for explaining firm behavior, as it is presented here. It is operationally identical to the standard theory of firm behavior and fails to explain the "perverse" relation between the dentist/population ratio and dental fees. While it may be apparent that one must return to the costless demand creation model to successfully explain the "perverse" response (and unsuccessfully explain virtually all other evidence), the following section presents a profit-maximizing model of firm behavior that explains the "perverse" results as a predictable consequence of efficient markets.

[7]See Robbins [14]. I must acknowledge discussions with Professor Ray Battalio of Texas A&M University who indicated to me the similarities between the targeted income behavior and Robbins' demand curve.

COMPETITIVE THEORY OF MARKET PERFORMANCE

The previous section argues that interest in targeted income behavior is based upon the single "perverse" result from cross-sectional studies (i.e., increases in the provider/population ratio promote higher fees). Recall that in these studies the provider/population ratio is considered an exogenous variable. Moreover, fees are assumed to record all components of price while empirically we find that they do not.[8] In the analysis below, we argue that in dentistry (as well as other care markets), the provider/population ratio reflects an equilibrium balance of market forces and cannot be regarded as exogenously determined. As an integral part of the analysis, patient waiting time is viewed as an additional component of price which registers the patient's optimal form of payment. Under quite plausible conditions, a perfectly competitive market promotes the "perverse" response. What has been regarded as the basis of interest in the targeted income hypothesis in fact supports conventional theory.

Consider a market such as dental care wherein the patient must be present at the time of service production. The consumption of the service, therefore, requires payments of time and money. We can express the patient's *full* price, P, as:

$$P = p + WT \tag{5}$$

where p represents the money price, W represents the patient's value of time, and T represents the amount of time spent in consuming the service. For simplicity, we can consider T as waiting time in the dentist's office.[9]

For a detailed exposition of the determination of waiting time in a market, we can turn to recent developments in queueing theory. The most fundamental presumption in the queueing literature is that of stochastic arrivals and service times. This presumption is particularly applicable in the dental sector, since two commonly acknowledged attributes of the market are that demand for care is based upon random incidences of disease and injury, and that the dentist can never perfectly estimate treatment time. No two mouths are identical, and response to treatment is never perfectly predictable. And all firms face the management difficulties of absence and tardiness among employees. Hence, there is a positive variance in the distribution of treatment time for the simplest procedures.

Given the randomness of demand, the dentist typically smooths the irregular arrivals of patients by maintaining appointment schedules, but schedules cannot preclude late arrivals, no-shows, walk-ins and emergencies. In short, the dentist can never eliminate the unpredictability of arrivals of patients. It is natural, therefore, to consider stochastic queueing models in

[8]See Saving *et al.* [15].

[9]In dentistry patients demanding non-emergency visits wait in the appointment queue as well. A discussion of the queueing process in dentistry and empirical results are presented in House [10] and Saving *et al.* [15].

explaining production and distribution decisions in the dental firm.

Among these stochastic models, we find that the literature relies largely upon the Poisson specification. This distribution has been shown to depict accurately arrival patterns in other studies and offers the convenience of a one-parameter distribution.[10] Its probability distribution is of the form:

$$P(K;\lambda) = e^{-\lambda}\left[\frac{\lambda^k}{k!}\right] \tag{6}$$

where λ is the single distribution parameter measuring both mean and variance. We apply this specification to both arrivals and service times, and assume that λ measures the expected patient arrival rate and μ measures the dentist's expected service rate. Although both follow the Poisson distribution, the two parameters, λ and μ, are not equal. The service rate is interpreted as the maximum rate of service delivery (one per patient). Of course, the maximum is maintained only when the dentist always has a patient to treat. During idle periods, the *actual* service rate falls below μ. We prohibit by assumption patient balks so the expected arrival rate, λ, is the *net* arrival rate (after balks).

It can be shown that with a Poisson arrival rate, the interval between patient arrivals follows an exponential distribution. Likewise, the interval between patients leaving the dentist's operatory follows an exponential distribution since the dentists service rate follows a Poisson distribution. By examining the exponential distribution closely, one discovers the Markovian or memoryless property of arrivals and departures. That is, the dentist does not "learn" to alter probable arrival times by past experience. For instance, the time that has passed since the last patient's arrival has nothing to do with the probability that the next patient will arrive within the next t minutes. This absence of learning ensures a degree of simplicity in the queueing model and perhaps accurately describes observable patient arrival patterns.

Consider the simplest queue discipline, first-in-first-out. With this discipline we ignore emergencies that break in front of the queues. Assume that the dentist maintains a "single-channel" in that each patient must be attended by the dentist at least once during treatment. If we mathematically examine the properties of such a queueing model, we find that for a stable solution (that is, waiting lines do not grow to infinite lengths), it is necessary that $\mu > \lambda$. The mean service rate must exceed the mean arrival rate. If this condition is satisfied, the patient's expected total wait (in the queue and in the operatory) is determined in the equation:

$$T = \frac{1}{\mu - \lambda} \tag{7}$$

where T is the expected wait. Notice that an increase in the expected service rate decreases the expected waiting time for the typical patient.

By combining equations (6) and (7), we can easily illustrate the relation between money price and the dentist's service rate *for a given full price.* That is:

$$\frac{\partial p}{\partial \mu} = W\left(\frac{1}{\mu - \lambda}\right)^2 \tag{8}$$

At the constant full prices, an increase in the dentist's service rate permits an increase in money fees.

The dentist may choose his service rate, μ, by employing more or less capital equipment and/or dental auxiliaries. The high speed handpiece as an example has long since been given credit for reducing treatment time. Recently treatment times have been reduced with the use of panoramic X-ray and more rapid pain prevention drugs. The use of auxiliaries clearly reduces treatment time, especially if trained in 4-handed dentistry. In short, there are techniques available to the dentist which significantly increase his expected service rate. In a simplistic form, consider a service rate function:

$$\mu = \mu\,(K,L) \tag{9}$$

where K and L measure the dentist's capital and labor inputs, respectively. Assume that the total cost curve associated with μ is "typical" in that it exhibits the convention (μ-shaped long-run average cost curve).

For simplicity, assume that the market for dental care is perfectly competitive.[11] Each firm attempts to maximize profits represented in the following profit function:

$$\pi = px - c(x,T) \tag{10}$$

where x represents total output and T records the expected waiting time for patients for a given patient arrival rate common to all firms. Each firm must establish the market-determined expected full price, P, although there exists an infinite set of money price-expected waiting time combinations that satisfy the full price constraint. For a constant output and zero economic profits, we can derive the competitive trade-off between p and T by examining the total differential of (10).

$$0 = x\,dp - \frac{\partial c}{\partial T}dT$$

$$\frac{\partial p}{\partial T} = \frac{1}{x}\frac{\partial c}{\partial T} < 0$$

$$\frac{\partial^2 p}{\partial T^2} = -\frac{1}{x}\,\frac{\partial^2 c}{\partial T^2} \gtreqless 0$$

Clearly, there exists an inverse competitive association between money price and expected waiting time. however, it is necessary to examine the second derivative to determine if there exists a unique competitive solution that establishes an equilibrium money price and expected waiting time combination among the infinite set. Using equations (7) and (9), one concludes that:

[10] This discussion is formally presented in Gross and Harris [8].

[11] Alternatively, we could assume that licensing restricts entry, and rents accrue to those existing firms in the industry. However, competition exists among firms such that rents are equally distributed.

$$\frac{\partial^2 p}{\partial T^2} > 0 \text{ if } \frac{\partial^2 c}{\partial \mu^2} \leq 0$$

$$\frac{\partial^2 p}{\partial T^2} \gtreqless 0 \text{ if } \frac{\partial^2 c}{\partial \mu^2} < 0$$

Hence, the competitive firm's trade-off between p and T is concave to the origin if the marginal cost of increasing the service rate either increases with μ or remains constant. Adopting either of these conditions, we can graphically present an iso-profit curve that illustrates the firm's money price-expected waiting time trade-off (see Figure 2).

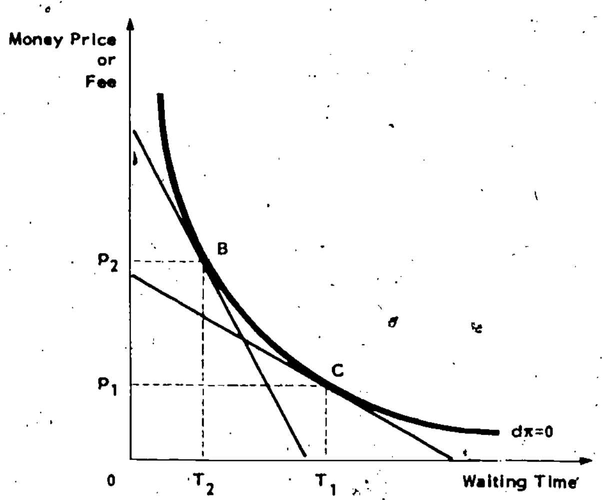


Figure 2

On the demand side of the market, consumers are assumed to search for the lowest full prices.[12] Consider a patient's utility function:

$$U^* = U^*(x^*, \ell^*, Z^*) \tag{11}$$

where x^* represents dental care, ℓ^* represents leisure time, and Z^* represents all other market goods. With the usual time constraint ($1 = L^* + \ell^* + T$ where L^* represents labor time) and income constraint ($L^*W = px + P_z Z^*$ where W represents the wage rate and p and P_z represents respective market prices), we can derive the utility maximization condition:

$$U_x = \sigma\{WT + p\} \tag{12}$$

Clearly, the consumer maintains a trade-off between p and T which is determined by the consumer's wage rate, W.

Consider two groups of consumers each exhibiting different values of time, W_1 and W_2. By illustrating each group's trade-off between p and T in Figure 2, we determing that two types of firms will exist in a competitive market. Members of group 1 with the lower value of time W_1 seek care from a firm offering the full price combination (p_1, T_1). Members of group 2 will seek care from a firm offering the full price (p_2, T_2).

In this market, the full prices paid by each group are approximately equal. Although the wage rates among members of each group differ, assume equal levels of demand for care among all individuals regardless of group membership. The aggregate arrival rate for each group, λg, is determined by the market full price. Individuals seek the lowest full price among all existing firms. Firms offering the (p_1, T_1) combination (type A firms numbering n_A) collectively expect a λ_A arrival rate where

$$\lambda_A = \lambda g \frac{n_A}{n_A + n_B} \tag{13}$$

Firms offering the (p_2, T_2) combination (type B firms numbering n_B) collectively expect a $_B$ arrival rate where

$$\lambda_B = \lambda g \frac{n_B}{n_A + n_B} \tag{14}$$

Each firm employs labor and capital to maximize profits. The cost function associated with varying service rates is assumed "typical" and is graphically presented in Figure 3. In equilibrium, each firm maximizes the Lagrangian function:

$$\pi = \lambda ph - C(\mu, h) + \delta\left[P - p - W\left(\frac{1}{\mu - \lambda}\right)\right] \tag{15}$$

where h represents the optimal length of "business hours" and δ is the Lagrangian multiplier. The optimal service rate must satisfy the condition:

$$\frac{\partial c}{\partial \mu} = \delta W\left(\frac{1}{\mu - \lambda}\right)^2 = \lambda h \frac{\partial p}{\partial \mu} \tag{16}$$

The equilibrium solution above reflects a balance between both cost and revenue conditions associated with μ. Increases in μ require increased expenditures toward factor inputs. Yet as μ increases, money price rises as expected waiting time falls. The competitive firm's optimal service rate, presented in Figure 3, is determined where the marginal revenue and marginal cost associated with changes in μ are equal. Note that the average revenue function is negatively sloped (due to changing price), and in competitive equilibrium the

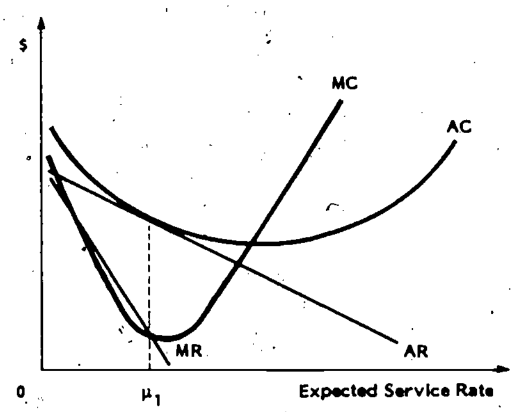


Figure 3

[12] Search behavior explains differences in prices paid by consumers. Those with a greater value of time search less and, on average, pay higher prices. The present analysis is unique in that time costs depend upon the firm's employment decisions and consumer choice and not upon search behavior of consumers alone.

average cost of μ is not minimized.[13] If a firm enjoys economic profits, competitors enter the market which reduces the expected arrival rate. Both AR and MR functions shift downward until the equilibrium tangency is maintained.

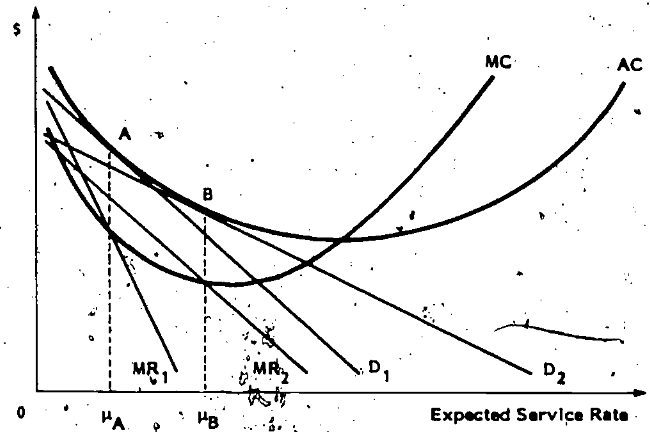


Figure 4

Assuming that all firms face the identical cost functions, we can examine the long-run competitive solutions for each type firm. With the Lagrangian function in equation (15), note that the properties of the total revenue function are constrained by the full-price equation. Marginal revenue is dependent upon both arrival rates and the value of the patient's time.

$$\frac{\partial[\lambda ph]}{\partial\mu} = \lambda hW\left(\frac{1}{\mu-\lambda}\right)^2 > 0$$

where

$$\frac{\partial^2[\lambda ph]}{\partial\mu\,\partial w} = \lambda h\left(\frac{1}{\mu-\lambda}\right)^2 > 0$$

where

$$\frac{\partial^2[\lambda ph]}{\partial\mu\,\partial\lambda} = hW\left(\frac{1}{\mu-\lambda}\right)^2 > 0$$

That is, the high income patients promote a unique demand curve facing the firm. The demand curve is *more elastic* than that promoted by low income patients. Again, arrival rates are altered with the entry of new firms and competition solutions are maintained. The equilibrium long-run service rates for the two type firms are presented graphically in Figure 4. As illustrated, a firm servicing the first group of patients (with lower values of time) optimally chooses a service rate, μ_A. Firms servicing the second group of patients optimally choose a greater service rate, μ_B.

In a competitive market with unrestricted entry of firms, we can now examine the equilibrium relationship between money price and firm density. Recall that empirically economists have observed a direct association between money price and dentists per capita. The analysis below illustrates the fact that this direct association between price and firm density is a probable result in a competitive market.

Consider two separate local markets each consisting of identical levels of demand for dental care. In both markets the aggregate arrival rates, λg, are equal and both include populations of equal size. The important distinction between the two markets lies in their preferred form of payment: one population preferes a more time-intensive form of payment for dental care. Dentists can freely migrate between markets, and our objective is to determine the equilibrium dentist/population ratio in each market and examine equilibrium fees, given a uniform full price for both populations. It is appropriate to infer that one demanding population places a greater value on its time than the other because of the differences in the preferred form of payment.

Consumer equilibrium requires equal full prices with differing combinations of money price and waiting time. Consistent with Figure 2, one population prefers to pay (p_1,T_1) While the other prefers (p_2,T_2). We must now determine the combinations of each firm's expected service rate, μ, and the competitive equilibrium arrival rate in each market. Since the aggregate arrival rate is known, we can derive the equilibrium number of dentists servicing each population.

From equation (7) we know that a locus of points exists in μ-λ space that is consistent with a unique waiting time. Expanding equation (7) by taking the total differential, we find that if $dT = 0$ and $\mu > \lambda$, $\frac{d\lambda}{d\mu} = 1$. Figure 5 graphically presents this analysis wherein the first quadrant illustrates the optimal service rates with differing forms of payment and the fourth quadrant presents the iso-waiting time curves in μ-λ space. The third quadrant maps the relationship between the firm's expected arrival rate and the required number of firms. Since the firm's expected arrival rate is determined by the ratio between a constant aggregate arrival rate and

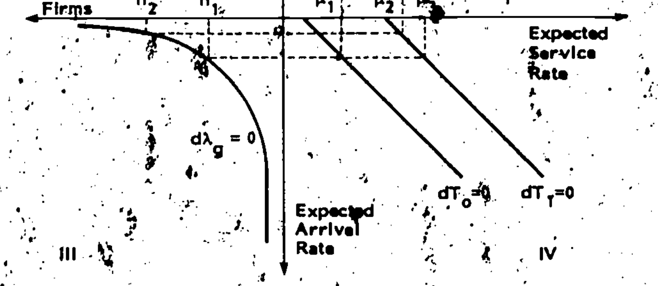


Figure 5

[13] Based on the changing quality of the care: as the service rate increases, waiting time is reduced reflecting improvement in one component of overall quality. For constant quality, the traditional tangency at the minimum average cost is ensured as output in the competitive firm is determined.

the number of firms. mapping is depicted by a rectangular hyperbola in n- λ space.

In the graphical analysis presented, the population preferring the less time-intensive form of payment faces an expected waiting time, T_1, and is serviced by dentists choosing an expected service rate, μ_2. This service rate is determined by the competitive equilibrium solution depicted by the tangency of the average cost curve and D_2. The other population faces a longer expected queue and dentists optimally choose an expected service rate, μ_1. It is clear that if full prices in one market were too high, excess profits would accrue to those dentists servicing this market and new firms would enter until full price returned to a competitive level. The solutions in Figure 5 illustrate only the equilibrium competitive solutions.[14]

From the expected waiting time function in equation (7), a unique optimal expected arrival rate exists in each market consistent with the expected waiting times and expected service rates. In quadrant III, these equilibrium expected arrival rates require unique numbers of firms, n_2 and n_1. Note that since the populations in the two markets are equal, differences in the numbers of firms reflect differences in the dentist/population ratio.

The conclusion from the graphical analysis is straightforward. The equilibrium solutions in the two markets require differing money prices and differing dentist/population ratios. In Figure 5 we discover that due to the shape of the firm's average cost curve, the market with the lower expected waiting time, T_1, attracts a larger supply of dentists, n_2. Hence, *the dentist/population ratios differ because of the differences in the preferred form of payment for care.* Only if dentists expanded their expected service rate non-optimally to μ_3 would the dentist/population ratios be equal (as are full prices). The crucial finding is that the market with the higher money fees (and lower expected waiting time) in competitive equilibrium supports the larger stock of dentists. Any cross-sectional study of these two markets will report that fees are directly related to dentist/population ratios — the basis for interest in supplier-induced demand and the targeted income hypothesis.

This is not to say that an inverse association between fees and dentist/population ratios is inconsistent with this theory. Indeed, if the range of increasing returns to scale in Figure 5 is sufficiently large, firms would expand their expected service rates beyond μ_3 leading to the opposite result — the market with higher fees attracts a smaller stock for firms. What is important is that this analysis demonstrates that the direct association between the dentist/population ratio is not inconsistent with competitive equilibrium and services no basis for rejecting the competitive model or accepting notions of demand creation or targeted income behavior.

[14] The dynamics of the adjustment processes are themselves informative but are not examined here. Descriptions of the adjustment processes are presented in Hooke [10].

SUMMARY AND IMPLICATIONS

In review of the literature, we find that the basis for interest in the target income hypothesis is the "perverse" empirical association between fees and provider/population ratios as reported in cross-sectional studies. According to the usual interpretation given to these results, one is led to conclude that as firms enter a local market, prices rise. The conventional profit-maximization model, at this point, is discarded and notions of demand creation are introduced and "controlled" with a targeted income constraint. Such has been the foundation and development of the target income hypothesis.

By examining notions of demand creation in the context of provider behavior and market equilibrium, we find that several specifications of demand creation activity are, on the surface, intuitively appealing. By emphasizing patient ignorance, one can argue that demand creation is costless to the provider. This, however, promotes successful explanation of the perverse association between fees and the provider/population ratio but fails to explain many other facets of market behavior such as provider migration patterns. With the introduction of a direct cost associated with demand creation, the provider is appropriately viewed as a profit-maximizing entrepreneur with a backward bending supply curve of labor — a theoretical construct that does not substantively differ from existing models of the profit-maximizing firm and the utility-maximizing entrepreneur. In short, the only specification of targeted income behavior that explains the perverse result likewise fails to explain most other aspects of market activity and firm behavior.

In dentistry we find that cross-sectional examinations of fees do not accurately record differences in full price. Patient waiting time serves a rationing function in local markets and represents a previously neglected form of patient payment. By combining elementary queueing theory with the profit-maximizing theory of the firm, one concludes that the direct association between fees and the dentist/population ratio is not at all perverse but a probable consequence of an efficient, competitive market solution. This theory seriously questions the basis for interest in the target income hypothesis.

While this paper focuses mainly on the target income hypothesis and the relation between fees and the provider/population ratio, the imporance of this analysis extends beyond these issues. The theoretical implications offer a different view of health care markets and the pricing mechanism. At least in dentistry, both fees and waiting times serve as market signals to suppliers and demanders. An equilibrium relationship exists between fees and waiting times, which is indicative of efficient market solutions. Empirical examinations of fees alone are not sufficient

measures of full price. Existing policy based upon such results are potentially unsupportable with examinations of full price. To the extent that other health care markets function similarly with market-determined money price-waiting time combinations, much of the existing evidence must be reexamined, as most current policy. Indeed, estimations of impacts of national health insurance, hospital closures, certificate-of-need legislation, and hospital rate regulation all potentially underestimate market response when based upon changes in money prices or total health care expenditures. A policy that forces patients to pay lower fees but face longer queues may successfully reduce inflation in the health care sector (as presently measured), but such reduction is not necessarily indicative of welfare gains among the population that policy is intended to serve.

References

1. Acton JP: *Demand for Health Care Among the Urban Poor, with Special Emphasis on the Role of Time.* Rand Corporation, R-1151-OEO, April 1973
2. Benham L, Maurizi A, Reder MW: Migration, location and remuneration of medical personnel: physicians and dentists. *Review of Economics and Statistics* 50:3:322-347, August 1968
3. Dyckman ZY: *A Study of Physicians' Fees.* Council on Wage and Price Stability, staff report. Washington, D.C., 1977
4. Evans RG: *Price Formation in the Market for Physician Services in Canada.* Ottawa, Queen's Printer, 1972
5. Evans RG, Parish EMA, Sully F: Medical productivity, scale effects, and demand creation. *Canadian Journal of Economics* 6:376-393, August 1973
6. Feldstein MS: The rising price of physicians' services. *Review of Economics and Statistics* 52:121-13[illegible], May 1970
7. Fuchs VR, Kramer MJ: *Determinants of Expenditures for Physicians' Services in the United States, 1948-68,* National Bureau of Economic Research, Occasional paper 117, DHEW Publication No. (HSM) 73-3013, 1972
8. Gross D, Harris CM: *Fundamentals of Queuing Theory.* New York, John Wiley & Sons, 1974
9. Holtmann AG, Olsen EO: The demand for dental care: a study of consumption and household production. *Journal of Human Resources,* 11:546-560, 1976
10. House D[illegible]: *The Economic Relationship Between Dentist's Income and Time Supplied.* Washington, D.C., DHEW contract (HRA) 231-77-0117, 1978
11. Kushman JR, Scheffler RM: Pricing health services: verification of a monopoly pricing model for dentistry *Journal of Human Resources* 13:402-415, 1978
12. Newhouse JP: A model of physician pricing. *Southern Economic Journal* 37:174-183, October 1970
13. Newhouse JP, Phelps CE: New estimates of price and income elasticities of medical care services. In: *The Role of Health Insurance in the Health Services Sector* (Rosett RN, ed.). New York, National Bureau of Economic Research, 1976
14. Robbins L: On the elasticity of demand for income in terms of effort. *Economica* 123-129, June 1930
15. Saving TR, Battalio RC, DeVany AS, Gramm WL, House DR, Kagel JH: *Labor Substitution and the Economics of the Delivery of Dental Services.* Washington, D.C., DHEW contract (HRA) 231-77-0135, 1978
16. Scheffler RM, Kushman JE: A production function for dental services: estimation and economic implications: *Southern Economic Journal* 44:25-35, 1977
17. Shepard L: Licensing restrictions and the cost of dental care. *Journal of Law and Economics* 21: 187-202, 1978.

Discussion

Frank A. Sloan, Ph.D.
Vanderbilt University

INTRODUCTION

Whether or not dentists and physicians can induce demand for their services and set income in accordance with a preconceived target has important implications for regulatory policy. If demand can be induced, it is desirable to have restrained utilization of medical services; if price differences reflect motivations of practitioners to attain an income target, and this factor is important, there is certainly reason for controls over price. From the vantage point of public policy, these are important issues worthy of serious study.

Several papers at this conference point to serious methodological flaws in some past work on physicians and dentists and suggest that the case of induced demand and target income-setting thas been overstated. The evidence, both theoretical and empirical, is not nearly as strong as proponents of inducement and target income-setting have asserted. This is not to say that physicians cannot induce demand under any circumstances. It may well occur, particularly in specialties with which individual patients have little contact over their lifetimes and for which third party reimbursement is relatively complete. The case for serious inducement in many primary care fields is probably much more limited. At present, empirical evidence is too scanty to lend support to these conjectures. In fact, most investigators have not even attempted to make such distinctions.

Four types of models of physician and dentist behavior have been discussed at this conference and in the literature in general. As seen in the rows of Figure 1, there are standard and non-standard theories and in the columns, "naive" and "sophisticated" schools within each category. In cell 1 is the standard profit-maximizing theory, which omits certain important features of health services markets but at the same time yields powerful predictions, e.g., price falls, *ceteris paribus*, as the supply curve (or the firm's marginal cost curve in the monopolistic case) shifts outward. Extensions of standard theory include the Pauly-Satterthwaite paper presented at this conference and quality-amenities model developed in several of my papers with co-authors.[1]

The Evans (II) model described in the Ramsey paper is an example of cell 3. As Ramsey states, the model is week conceptually and not really consistent with observed facts. More promising is an extension, Evans (1), which assumes a preference function for the practitioner, containing profits, workload, and discretionary power as arguments. Unfortunately, the model yields few refutable hypotheses because of offsetting effects.[2] In a way, Evans (1) is similar to advertising models which economists have been working with for years. In this sense, even Evans (1) could be considered a "standard" economic model of the firm.

Much of the empirical literature on this topic contains a number of non-trivial deficiencies. One hopes that many of these will be corrected in future work. Meanwhile, we economists should be quite cautious about making strong statements that we have firm evidence that markets for dentists and physicians are anomalous when we don't.

First, some regressions are seriously misspecified or under-specified. One study of physician fee-setting that contains one explanatory variable has been cited repeatedly in the debate on the effects of physician and dentist availability on fees. Often fees are expressed in nominal rather than in real terms (corrected for differences in the area cost-of-living). A table in a recent study by the U.S. Council on Wage Stability shows how sensitive empirical results can be, especially on the estimated coefficient of the physician-population ratio

[1] For a review see Sloan and Bentkover [6].

[2] In addition to the Ramsey paper, see Sloan and Feldman [7].

variable in a fee regression depending on whether or not the dependent fee variable has been deflated. Significantly positive physician-population ratio coefficients became essentially zero.[3] Among key omitted variables is health insurance coverage; this variable should be measured more precisely and comprehensively than merely the percentage of the population covered, e.g., measures of depth of coverage, methods of third party payment.

	"naive"	"sophisticated"
Standard theories	1	2
Non-standard theories	3	4

Figure 1. Alternative Models of Practitioner Behavior

Second, authors have frequently made internally inconsistent inferences from their empirical evidence. For example, they are quick to point out that a positive coefficient on a physician-population ratio variable in a fee regression is evidence for target income-setting and/or inducement, but at the same time, positive coefficients on other pertinent variables, such as patient income, which are fully consistent with standard economic theory are not emphasized. A finding of positive physician-population ratio coefficients *and* positive income coefficients would seem at a minimum to yield a standoff between standard and non-standard models.

Third, sometimes the underlying theory has been disregarded as useless because it yields ambiguous predictions; but once empirical results are presented, authors make strong statements about the relationship of theory to empirical results obtained. If a variable X is predicted from theory to have a positive *or* negative effect on Y, it is clearly inappropriate to say that a positive or negative coefficient, obtained in the empirical phase of the analysis, supports the theory.[4]

Fourth, most studies of dentists and physician behavior do not consider inter-practice variations in quality-amenities. The assumption that output is homogenous is much too strong and leads one, for example, to attribute a price difference wholly to (a) market imperfections and anomalies, including target income-setting, or (b) temporary or even persistent disequilibrium. Price differentials may exist in equilibrium because of variations in quality-amenities (such as waiting time in the physician's office, length-of-visit). Providers tend to stress quality matters in the policy debates. In this instance, their case is stronger than many analysts' cases (even though providers' statements may also be self-serving). Quality in this context should be defined much more broadly than in strictly medical terms. Included are various dimensions of patient access to health care services.

Fifth, frequently insufficient attention has been devoted to the definition of the appropriate market area. If the market area is defined inappropriately, a border-crossing problem may result, which can in turn lead to serious biases in the estimated parameters and misleading policy inferences from the empirical results. The Held-Manheim study is to be commended for its attention to this issue.

Sixth, fees and dentist or physician-population ratio may be simultaneously determined. This simultaneous relationship may partly explain the observed positive coefficients on physician-population variables in fee regressions.[5]

Seventh, insufficient attention has been devoted to occupational differences, in this context physicians and dentists, and variations among specialists within a particular field, especially important for medicine. Frequently evidence on doctors is used uncritically in studies of dentist behavior, even though their goals and constraints could be quite different, as the Littleton paper presented at this conference stresses. Further, as already noted, it is possible that inducement is far more important in some physician fields than in others.

IMPLICATIONS FOR FURTHER RESEARCH

Compared to two or so years ago, conceptual and empirical research on the subject of this conference has made important strides. Many now hold a much more cautious view of supplier-induced demand and the target income hypothesis, even though weakness in standard theory in this context are also recognized. New, comprehensive data bases are now available. The data base used by Held-Manheim is only one of these. Because of recent progress, there is greater reason to be confident that criticism can serve a constructive purpose. We can be more critical about the way specific variables are constructed because large-scale surveys are being conducted and refinements can be incorporated in these surveys.

This subject commands sufficient interest that conceptual or empirical developments should have a ready-made audience, both in the academic world and in the public policy arena. This interest should serve to stimulate new research and government and private financial support for this work.

Finally, although this conference has largely been made up of economists, certain aspects of future

[3]See U.S. Council on Wage and Price Stability [8]. Dr. Zachary Dyckman, the author of this report, stated at this conference that the physician-population ratio coefficient became positive again, once regional dummy variables were removed. These regressions were not presented in the report. There is some question whether the regional variables belong in his fee regressions; it would have been far preferable to include key area variables instead, e.g., measures of health insurance. But without these area variables, it is probably better to include the regional variables as a second-best solution.

[4]Evans [1] is subject to this criticism.

[5]See Frech and Ginsburg [3] for further discussion of this point.

research will require interdisciplinary efforts, especially between dentists and physicians on the one hand and economists on the other. This joint research can include further probing into institutional features of these markets, a better definition of output as well as the choice of tracer illnesses.

References

1. Evans RG: Supplier-induced demand: some empirical evidence and implications. In: *The Economics of Health and Medical Care* (Perlman M, ed.). London, Macmillan, 1974, pp. 162-173
2. Feldstein MS: The rising price of physicians' services. *Review of Economics and Statistics* LII:124-133, 1970
3. Frech HE, Ginsburg PE: Physician pricing: monopolistic or competitive: comment. *Southern Economic Journal* 38:573-577, 1972
4. Fuchs VR: The supply of surgeons and the demand for operations. *Journal of Human Resources,* forthcoming 1978
5. Fuchs VR, Kramer MJ: *Determinants of Expenditures for Physicians' Services in the United States, 1948-68.* Washington: National Center for Health Services Research and Development, December 1973
6. Sloan FA, Bentkover JD: *Access to Ambulatory Care and the U.S. Economy.* Lexington, MA, Lexington Books, forthcoming 1978
7. Sloan FA, Feldman RD: Competition among physicians. In: *Competition in the Health Care Sector: Past, Present and Future* (Greenberg W, ed.) Washington, Federal Trade Commission, 1978, pp. 57-131
8. U.S. Council on Wage and Price Stability: *Physicians: A Study of Physicians' Fees.* Washington, U.S. Government Printing Office, March 1978.

Discussion

Zachary Y. Dyckman, Ph.D.
Council on Wage and Price Stability

We have heard much criticism of the target income hypothesis as applied to both dentists and physicians and not a word in support of it. Unfortunately, Mark Evans, the leading proponent of the hypothesis, could not attend the workshop. I assume I was chosen in his place because I have been associated, somewhat incorrectly, with the target income hypothesis. I am not prepared to defend Evans' formal model of target income behavior, although I do believe the hypothesis has some validity in explaining the economic behavior of some groups of physicians.

My comments are addressed principally to Professor Ramsey's paper which deals formally and almost exclusively with the target income-standard market model debate. The target income hypothesis has few supporters. Physicians do not like it because it implies, at least to them, that physicians are greedy. They just decide how much they should be earning and adjust their fees and, perhaps, quantity or service mix accordingly. Economists do not like it because it suggests that the price and quantity of physicians' services are not determined primarily by the interaction of supply and demand forces, and that physicians are not profit maximizers. To oversimplify, physicians are distressed that they are portrayed as having enormous power over fees (and quantity) and economists are distressed because physicians fail to use all the market power they possess. Therefore, anyone supporting the hypothesis runs the risk of being violently drummed out of the profession and then having no one to minister to his wounds.

Ramsey, in a well written and almost elegantly logical paper, asks us to choose between the neoclassical or what he calls the standard model of the market for physicians' services and models of physician-induced demand and target income. Unfortunately he examines not the veracity of the theory of physician-induced demand and target-income pricing as it applies to physician behavior but the logical consistency of two specific theoretical models. The models are rejected in favor of the traditional market model, because of what he regards as inconsistency within the models and what he labels external inconsistency, or aspects of the hypothesis that are inconsistent with aspects of existing theory that are retained.

Professor Ramsey has made a positive contribution to the literature on models of physician behavior, but not, I believe, to determining the validity of the target income hypothesis. His contribution is in specifying requirements for logical consistency — both internal and external — and in describing methodological problems that plague most of the empirical work in this field, particularly that of specification error.

In his oral presentation, Ramsey makes it clear that for both theoretical and empirical reasons, the target income hypothesis merits rejection. In the written version, Ramsey is somewhat more circumspect. In fact, early in the paper he notes, "Unfortunately, as will be shown below there are no crucial tests to discriminate between the competing hypotheses considered in this paper." Later on, however, implicitly taking the view that a challenger (target income) must decisively beat the current champion (standard model) in order to dethrone it, he does reject one version of the target income hypothesis, although it appears more for lack of aesthetic properties of the model than for incompatibility with available evidence.

Ramsey considers two submodels of the target income hypothesis separately: (1) physician-induced demand and (2) target-income pricing. Ramsey has difficulty remaining impartial regarding the validity of these hypotheses relative to the standard market model — of price equilibrating independent supply and demand schedules. For instance, Ramsey doubts that consumers are less knowledgeable about medical care than other products and services. He speculates, "If consumers are ignorant, then risk aversion will lead them to engage in costly activities to alleviate that ignorance by checking medical opinions with other doctors, by reading medial books and popularized versions of medical books. . . ." I think Ramsey is missing the point. It may not be ignorance, but reliance on and trust in the physician's superior knowledge and expertise, particularly for seriously ill and hospitalized patients. Of course, patients' concern about the cost ramifications of physicians' decisions is reduced by their paying directly, on average, only 6 percent of the hospital bill and, for nonprimary care, 20-30 percent of the physician's bill.

Ramsey rejects the induced demand model, primarily

because "it provides fewer refutable hypotheses" and because "of the non-observability of D," where D is the effort required of the physician to induce additional demand for his services. In essence, the model is being rejected because, regardless of its conformance or non-conformance to real world phenomena, it is inconvenient to test empirically, and because of the lack of empirical evidence of time being spent to convince patients to purchase additional medical care, a factor that Ramsey assumes is integral to the induced-demand model. In fact, it is not clear that a physician providing excess care does spend more time discussing course of treatment than a conscientious physician providing what he regards as medically optimal level of care. Ramsey also disregards evidence, some of which cannot be neatly inserted in econometric models, of physicians ordering more diganostic tests in situations where they profit from those tests than where they do not.[1]

Actually, the concept of physician-induced demand is fully consistent with a profit-maximization model. What better way is there to increase profits than shifting the demand curve? It is only the *ability* of the physician to do so that irks the traditional economist, because it implies that market forces cannot be relied upon to achieve a desirable allocation of resources, or the economist's Nirvana, Pareto Optimality. Mark Blumberg has argued that physicians control 70 percent of total medical care expenditures and 93 percent of hospital expenditures.[2] One need not go that far to appreciate that physicians do wield substantial power over both the nature and quantity of care provided. Perhaps the best evidence of physician-induced demand is the physician response to the malpractice problem. A 1977 American Medical Association survey reported that 76 percent of physicians were ordering extra tests and procedures as protection against malpractice claims.[3] I think it is clear that in most cases the patient does not demand these tests.

Professor Ramsey may be on firmer ground when he rejects Evans' attempt at a formal theoretical structure for the target income hypothesis, because of some logical inconsistency internal to the model. Much of the inconsistency problem relates to the target income model being burdened with profit-maximization requirements. However, a basic precept of target income behavior is that some physicians charge less than a profit-maximizing fee, independent of hours of work, number of interesting cases and other elements in the physician's utility function.

The standard model is also preferred, according to Ramsey, because it is equally compatible with the often observed positive correlation between fee levels and physician/population ratios, during *a short-term adjustment path*. But it is folly to think that the observed correlations reflect only short-term adjustment periods, after which an increase in supply of physicians will lower fees. It remains clear that target income behavior is more compatible in a comparative status sense with a positive correlation between fees and relative physician supply than the standard market model. Also, as Reinhardt has shown, target income behavior is also compatible with a negative relationship between fees and physician supply.[4]

My own view of the market for physicians' services is that a conventional market model may be most useful for describing economic behavior for services of some specialists while a "non-market" model best describes economic behavior for others. However, within any market, defined geographically and by nature of physician service, both market and "non-market" factors combine to determine price and quantity. For some physician services, with the nature of service probably best defined by specialty of physician, market forces will dominate and for others non-market forces will dominate.

What is being suggested is that for the services of many physicians, demand is likely to be inelastic and market forces are not an important determinant of price. Howe then are prices determined? Is a target income chosen and prices selected to achieve that level of income? The accepted wisdom in the economic profession is that there is a strong tendency for profit maximization in most markets. In a competitive situation, the market enforces discipline and those that do not maximize profits sustain losses and eventually are forced out. Among physicians, however, with median earnings in private practice currently around $75,000, those that do not price to maximize profits can survive quite adequately, in fact better than most economists.[5] While extensive insurance coverage and other institutional factors enable most physicians to earn exceptionally high levels of income, they also afford the physician the luxury of not maximizing profits — by charging less than prevailing fees or not minimizing practice expenses — and still have higher earnings than those in other professions. The existence of economic rent for physicians, as demonstrated by Sloan and obvious to anyone without blinders, is a primary requisite for widespread target income pricing and other nonprofit optimization behavior.

To determine the relative importance of market forces and non-market factors such as target-income pricing, it may be useful to divide physician specialists into three categories: 1) hospital-based specialists (i.e.,

[1] Uwe Reinhardt, "Health Manpower Policy in the United States," Paper presented to the Bicentennial Conference on Health, Philadelphia, November 11-12, 1976.

[2] Mark S. Blumberg, "Rational Provider Prices — an Incentive for Improved Health Delivery," in George K. Chacko, Editor, *Health Handbook — 1978*, Amsterdam, North Holland Publishing Co., forthcoming.

[3] *AMA News*, March 28, 1977.

[4] Uwe Reinhardt, "Parkinson's Law and the Demand for Physician's Services," Paper presented to the *Conference on Competition in the Health Care Sector*, June 1 and 2, 1977.

[5] The $75,000 estimate is extrapolated from $63,000 earnings in 1976. Zachary Dyckman, *A Study of Physicians' Fees*, Council on Wage and Price Stability, 1978, p. 75.

pathologists and radiologists), 2) surgeons and some highly specialized internists, and 3) primary care physicians (including psychiatrists). In terms of factors that determine price, and to a lesser degree quantity or nature of services, we are talking about virtually three distinct markets.

Revenues of hospital-based physicians are in most cases tied to reimbursement of hospitals and, for many Blue Cross, Medicare and Medicaid patients, may be unrelated to fees listed on the bill provided to the patient. Another relevant factor is that more than 75 percent of physician revenue is derived from insurance sources (including Medicare and Medicaid). Also, it is usually other physicians who order services from radiologists and pathologists, although these physicians have some indirect influence on demand.

Fees charged by hospitals for radiology and pathology vary widely among hospitals and between hospitals and outside laboratories and offices. Given hospital reimbursement methods (radiology and pathology are generally paid for through the hospital), the lack of competition within a hospital for radiology and pathology, and the high degree of insurance coverage for these services, supply and demand forces can be expected to play only a very limited role in determining fee levels for radiology and pathology. More important are factors such as negotiating skill and influence of the physician within the hospital community, historical factors, involvement of third party payers in hospital charge structure and other institutional factors. I doubt that target-income pricing is very important because the non-salaried pathologist or radiologist typically earns over $125,000 per year, which is above the income of most other professionals and physicians.[6]

Surgeons, anesthesiologists and some highly specialized internists also receive most of their revenue — 70-85 percent — from insurance sources.[7] The dominant method of insurer payment for surgeons' services is the usual, customary, and reasonable (UCR) fee approach. The UCR method typically pays 80-100 percent of the fee provided it does not exceed the physician's usual fee or the 90th percentile fee in the area.[8] This payment approach insulates the physician from the constraints of normal supply and demand forces and gives him and his peers great latitude in increasing fee levels. In other words, the small portion of the fee usually paid directly by the patient and the generous UCR payment approach combine to make it likely that demand for surgery is highly inelastic.

Historical and institutional factors, in my opinion, tend to dominate normal market forces as determinants of surgical fee levels. Surgical fees vary substantially across areas exhibiting strong patterns by region and city size and a positive association with relative physician supply.[9] Surgeons' incomes tend to be relatively invariant across areas, with an apparent tendency for high fees to compensate for low volume per surgeon. This, in combination with what appears to be inelastic demand conditions, strongly suggests that physicians can (to a significant extent) and do increase fees to compensate for low volume in order to achieve some target level of income.

Primary care physicians (including psychiatrists) typically derive from one-fourth to one-half of their revenue from insurance sources.[10] They also practice in a more competitive environment, partly because of the substantial direct consumer payment for their services but also because these physicians typically provide more recurring services in a less crisis type atmosphere than surgeons and other specialists. This facilitates greater quality and price comparisons than for physicians rendering one-time care, often in more crisis type environments, or for physicians whom the patient may not select or come in contact with before services are provided. Given the stronger consumer incentive to minimize the direct cost of care because of lower insurance coverage for primary and psychiatric care, one would expect a greater role for market forces in price and quantity determination. Inelastic demand and pricing below the profit-maximizing level may be common for primary care physicians, but market forces should play a greater role in determining fee patterns for them than for surgeons and hospital-based physicians.

Dentist fees, because of relatively limited dental insurance coverage, are likely to be primarily determined by the interaction of normal supply and demand forces rather than non-market factors. However, because of the relatively high earnings level of dentists (relative to earnings for most occupations except physicians), pricing below profit-maximizing levels and target-income pricing may not be uncommon.

The question of physician-induced demand is not whether it occurs but how much of it occurs. As already indicated, I think the fact that most physicians admit ordering additional tests in response to the malpractice threat is an excellent indication of the power of physicians to induce demand. How much exists is difficult to determine. In view of the variation in medical opinions about desirable levels of care, even for specific medical conditions, it is usually difficult to distinguish from cross-section utilization or expenditure data whether regional differences reflect 1) underutilization in some areas related to inadequate supply or 2) excess utilization in some areas related to excess supply and induced demand. Conceivably, micro data at the individual physician level, analyzed both from an econometric and medical perspective, could help resolve this issue.

[6] *Ibid.*, pp. 79-81.

[7] *Ibid.*, p. 34. The 70-85 percent estimate is an extrapolation which reflects the 5-8 percent increase since 1970 in the proportion of physician revenues paid for by third party payers.

[8] *Ibid.*, pp. 23-33.

[9] *Ibid.*, pp. 111-127.

[10] *Ibid.*, p. 34. See footnote 7.

Discussion

Thomas R. Saving, Ph.D.
Texas A&M University

In the process of reading the papers for the conference, I was again struck by the same feeling of uneasiness I had when I first became interested in the economics of health care some eight years ago. I have traced this uneasiness to the lack of a clear relation between the somewhat anomalous empirical results and the theory to which these results are supposed to offer evidence. Several of the papers presented here have clarified my thinking on this issue and helped me to appreciate the general complexity of the problem.

Essentially, the problem lies in the related empirical results that fees per visit and the number of visits per illness for both physicians and dentists appear to be positively related to the density of physicians or dentists. At first blush one would suspect that such an anomalous result must be the result of a classic identification problem. That is, once one properly set up the empirical problem the demand and supply equation would be specified and an increase in supply, i.e., density of practitioners, would reduce price. It cannot be said, however, that the empirical work of Feldstein or others is obviously guilty of a naive econometric approach. Thus, the thrust of the work has shifted from working for an econometric reason for the anomaly to seeking a theoretical reason for it.

The principal theoretical twist chosen is that physicians or dentists deal in a product of which the consumers have considerable ignorance. Thus it is argued that both dentists and physicians can capitalize on this ignorance and affect demand. Given this assumption, it remains only to explain why the intensity of demand creation varies positively with the density of practitioners. Enter the existence of a target income and the anomaly is explained. Unfortunately, we are still hard pressed to explain the level of the income target and until we have a theory for it, we cannot predict the directional effect of many policies.

THE CONFERENCE

Having set the stage for the current conference, let me briefly discuss and possibly assess the contribution of the work reported on here. The papers can be split into two groups: those that attack the target income hypothesis on the empirical level and those that attack it on the theoretical level. In the first group are the papers by Held and Manheim and Hixson and Mocniak. The second group consist of papers by House, Ramsey, and Reohrig. The paper by Pauly and Satterthwaite spans both of these approaches and the Littleton paper is purely descriptive but nonetheless useful.

The empirical work reported consists of both cross-section and time-series estimation. Interestingly, the only time-series work (Hixson and Mocniak) results in demand and supply equations that are quite consistent with traditional economic theory. The cross-section work, on the other hand, is consistent with a positive relation between fees and practitioner density and thus with the target income hypothesis. It is important to note, however, that only the Pauly-Satterthwaite paper develops an explicit alternative to the target income theory and attempts to test the alternative hypotheses.

The three theoretical papers all attempt to develop internally consistent models of the target income hypothesis for the purpose of finding the set of prediction statements from these models. The essential question is whether the intersection of the prediction statements from the two theories contains as a subset the set of prediction statements from one or both theories. If this is true, then no empirical evidence can distinguish the two theories.

To distinguish one theory from another, the set of prediction statements of one theory must not be a subset of the other theory's prediction statements. With this in mind, let me discuss the set of theoretical papers. The Ramsey paper accomplishes, I believe, two things.

First, it develops internally consistent versions of a target income model. This development is a significant contribution since internal consistency is necessary before the set of prediction statements can be derived. Ramsey, in developing his version of an internally consistent target income model, assumes that a competitive market exists. This is crucial for his subsequent derivation of the set of prediction statements since it implies that profit maximization and utility maximization will be consistent goals. His conclusion that the target income set of predictive statements is a subset of the profit maximization set of prediction statements may not hold for a non-competitive market.

Significantly, the Pauly-Satterthwaite paper develops just such a non-competitive model. Their model, however, is one of the target income variety but is rather a traditional monopolistic competition model of the physician and dental markets. The focal point of the work is on the information system utilized in the physician-dentist market: that advertising is, in general, not permitted so that word of mouth is the principal information source. I find that I must take issue with a fundamental proposition of their increasing monopoly model. This proposition is that the level of information declines as the number of physicians rises. In particular they argue that given N physicians and M sources of information, the representative consumer has M/N units of information about each physician. Thus, the expected price differential required to make a move from one physician to another expected utility-maximizing physician will be greater the greater the number of physicians. If one assumes instead that the representative consumer has one unit of information about M physicians, then the number of physicians will have no effect on the price-differential required for an expected utility-maximizing move. Additionally, the problem at hand is one of Knightian uncertainty rather than risk and it is well known that the expected utility hypothesis does not organize choice consistently under Knightian uncertainty. See, for example, the work on the "Ellsberg Paradox."

The Roehrig paper is similar to the Ramsey paper in that it is an attempt to develop an internally consistent version of the target income model. It is distinguished from Ramsey in that it is based on a version of a monopoly model. This aspect of the paper is asserted rather than derived and I suggest that the type of non-competitive model assumed must be formally introduced into the structure of the model before the model can be tested for internal consistency.

House's paper is devoted to developing a non-target income model which is nonetheless consistent with the anomalous empirical results which have led to this conference. The paper represents an excellent discussion of the problems with developing an internally consistent version of the target income model. Interestingly, if "a la Baumol" target income behavior is treated as a "rule of thumb" similar to target pricing by firms, the anomaly would follow in the short run but not in the long run. That is, during the initial phase of an exogenous shift outward in supply, prices would rise but the market would force a long-run downward adjustment in price.

SOME GENERAL REMARKS

Having briefly discussed the papers presented, let me close by commenting on the problem generally. I am reminded of the early debate between the monetarists and the fiscalists. In that debate, great amounts of effort were expended on developing empirical evidence that turned out to be irrelevant for the purposes of settling the debate. That is, it became a question of what evidence was no evidence for which theory. That the evidence presented neither proved nor disproved a theory did not imply that such empirical information was of no value. I believe the current conference is a case in point. The evidence presented up to this time is information which will be useful in ultimately developing the underlying theoretical construct that organizes the information. The theoretical papers presented here represent an important step in that direction.

On the theoretical level, I suggest that we must recognize that physician visits involve considerable time and travel cost and that the quality of services varies. Thus, it is perfectly consistent for the increases in fees associated with increased density of physicians to be associated with a reduced full-cost of constant quality care. That is, waiting time is reduced, physicians are idle a greater proportion of the time and the fee plus the time cost of care is lower than when the physician density is lower. In a sense this is demand creation in a manner similar to the existence of inventories which reduce waiting time and therefore increase demand.

What the above suggested model emphasizes is that both prices and queues ration. In some markets, price is both the principal short- and long-run rationing device, e.g., the stock exchange, grain exchange. In other markets, queues are the principal short-run rationing device, e.g., restaurants. A significant, but as yet underdeveloped contribution will be had when someone develops a theory that organizes the various markets according to the extent that prices and queues ration. I believe that such a model will be an important cog in our understanding of the medical and dental markets. I believe it will relieve us of being in the position of having a special theory for the medical and dental industries.

BHM
·S ·DEPARTMENT OF
EALTH, EDUCATION, AND WELFARE
ublic Health Service
ureau of Health Manpower
HEW Publication No. (HRA) 80-27 January 1980